Version 2 CLIMCAPS-*Aqua* Retrieval Product Performance Test Report

Edited by:

Qing Yue, Bjorn Lambrigtsen, Tao Wang, and Jacola Roman

With Contributions by:

Eric J. Fetzer, Evan Fishbein, Fredrick W. Irion, Brian H. Kahn, Peter Kalmus, Evan Manning, Ruth Monarrez, Thomas Pagano, Joao Teixeira, Yixin Wen, Sun Wong

January, 2021

Jet Propulsion Laboratory, California Institute of Technology

Pasadena, CA

Submit Questions to:
https://airs.jpl.nasa.gov/data/support/ask-airs/

Table of Contents

The research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (80NM0018D0004).

1. Technical Summary

The Community Long-term Infrared Microwave Coupled Atmospheric Product System (CLIMCAPS) was developed to retrieve multiple parameters to describe the atmospheric, cloud, and surface states from a multi-decadal hyperspectral infrared and microwave observational record of AIRS on *Aqua* (since 2002), and CrIMSS on *SNPP* (since 2011) and the *JPSS* series (since 2017). The algorithm uses an Optimal Estimation methodology in which the Modern-Era Retrospective Analysis for Research and Applications version 2 (MERRA2) interpolated to both time and space of the satellite observations is used as an a-priori first guess.

This report presents the initial results of analyses performed by the JPL AIRS Project to assess the general quality of various core products produced by the Version 2 (V2) CLIMCAPS-*Aqua* retrieval system. Initial science assessment is performed by quantification of the retrieval yield, retrieval biases and root mean square errors, and the sampling biases in different regions and physical conditions. The objective was to evaluate the general quality of the retrieval by measuring whether the CLIMCAPS-*Aqua* V2 core products are within reasonable range of well-established reference data sources including in-situ and ground observations, other well-validated satellite observations, and reanalysis. For a complete description of the CLIMCAPS algorithm, please refer to the CLIMCAPS Level-2 ATBD (Smith and Barnet 2019, 2020). For detailed descriptions on various reference datasets used in the testing, readers are also advised to refer to Yue and Lambrigtsen (2019) for AIRS Version 7 Level 2 Performance Test and Validation Report and Wang et al. (2019) for the testing report on CLIMCAPS Level 2 Water Vapor and Temperature Vertical Profiles by the CrIMSS instruments on *SNPP* and *JPSS*-1.

This report provides analyses on the following categories of core V2 CLIMCAPS-*Aqua* retrieval variables:

- Thermodynamic quantities
    - Profiles of temperature and humidity
    - Total column water vapor
    - Near surface properties
    - Cloud parameters
- Trace gases such as ozone

Complete lists of all variables from CLIMCAPS can be found in the Level-2 Products User Guide: File Format and Definition (Monarrez et al. 2020).

CLIMCAPS-*Aqua* produces two different retrievals: the IR+MW (AIRS/AMSU infrared and microwave combined) and IR-only (AIRS-only). The IR+MW retrieval temporal coverage stops on September 24, 2016 due to the complete loss of AMSU-A2, while the IR-only products provide a data record covering the entire AIRS operation period. Both retrievals are evaluated in this report and the products are found to have no significant differences from each other.

2. Statistical Metrics Used in the Evaluation

The tested version is CLIMCAPS-*Aqua* V2.39, which is the official V2 release of the retrieval. The retrievals are evaluated using four metrics: the retrieval yield, retrieval bias, root mean square errors or differences (RMSE), and the sampling bias, which will be elaborated below. The L2 retrieval quality control (QC) indicators are used. The name of each QC indicator is the same as the name of the corresponding parameter with "_qc" appended to the name. For CLIMCAPS, QC=0 or 1 indicates the data products individually meet the accuracy requirements, and QC=2 indicates the use of such data is not recommended. Level 3 (L3) products are produced using L2 data with QC=0 or 1.

Below are the statistical metrics used in the quantitative evaluation of CLIMCAP-*Aqua* retrieval products in this report:

- **Retrieval yield** profiles are calculated as the percent of successful retrievals (passing the QC) at each level/layer.

- **Retrieval biases and RMSE** are calculated using equations below where both the truth and the retrievals are filtered with the L2 QC indicators.

  For temperature:

$$bias \equiv mean[x_{data}(QC) - x_{truth}(QC)] \tag{2.1}$$

$$RMSE \equiv \sqrt{mean\{[x_{data}(QC) - x_{truth}(QC)]^2\}} \tag{2.2}$$

  For water vapor:

$$bias \equiv mean[x_{data}(QC) - x_{truth}(QC)]/mean[x_{truth}(QC)] \times 100\% \tag{2.3}$$

$$RMSE \equiv \sqrt{mean\{[x_{data}(QC) - x_{truth}(QC)]^2\}}/mean[x_{truth}(QC)] \times 100\% \tag{2.4}$$

- **Sampling biases** are caused by the cloud-state-dependent sampling of the infrared sounding instrument (Fetzer et al, 2006; Yue et al. 2013, Tian et al. 2020) and the sensitivity to the surface conditions (Yue and Lambrigtsen, 2017). It is defined as the following equations.

  For temperature:

$$Sampling\_bias \equiv mean[x_{truth}(QC)] - mean(x_{truth}) \tag{2.5}$$

  For water vapor:

$$Sampling\_bias \equiv \{mean[x_{truth}(QC)] - mean(x_{truth})\}/mean(x_{truth}) \times 100\% \tag{2.6}$$

- **Skill score** is a unitless number showing the reduction of mean square differences or errors (**MSE**s) of the retrievals with respect to a given target dataset, in which the MSEs of both the retrieval and the target data are calculated using the same reference dataset.

$$Skill\ Score = 1 - \frac{MSE_{retrieval}}{MSE_{target}} = 1 - \frac{mean((T_{retrieval} - T_{reference\_data})^2)}{mean((T_{target} - T_{reference\_data})^2)} \tag{2.7}$$

  Positive (negative) skill score indicates retrievals have smaller (larger) MSE than the target data, thus closer to the "truth" represented by the reference data.

3. Summary of Version 2 CLIMCAPS-*Aqua* Retrieval Performance Testing Study

Based on the current analyses using the statistical metrics listed in Sec.2, following major findings are found for the retrieval performance of the V2 CLIMCAPS-*Aqua* products evaluated in this study:

- The performance is very similar between the CLIMCAPS-*Aqua* IR+MW and IR-only retrievals.
- The CLIMCAPS-*Aqua* yield has much smaller vertical variation below 250hPa compared to AIRS V7, reflecting a more "whole profile" quality control in CLIMCAPS-*Aqua* producing a more constant yield throughout the profile. The yield is smaller in CLIMCAPS-*Aqua* than AIRS Version 7 (V7) in the upper troposphere for partial cloudy conditions but much larger between 700 hPa and the surface. The yield in general decreases as effective cloud fraction (ECF) increases because of the impact from clouds on infrared retrievals. However, CLIMCAPS-*Aqua* also shows a decrease of yield over land with ECF < 0.2 especially over snow- and ice-covered surfaces, which is likely caused by the heterogeneity of land and frozen surfaces that increases the uncertainty in the cloud clearing.
- CLIMCAPS-*Aqua* produces a global dataset of temperature and water vapor vertical profiles with quality comparable to the AIRS V7 products. The mean bias magnitudes are within ±10% in water vapor vertical profiles and within ±1K in temperature. CLIMCAPS-*Aqua* produces a much smaller RMSE than AIRS V7 when comparing with both reanalysis and in-situ observations and the error estimate reported in the retrieval product is a more realistic indicator of the retrieval uncertainty. This indicates that the CLIMCAPS-*Aqua* retrieval performance is stable across different scenarios, which is partially related to the more stable a priori, MERRA2, used in the retrieval.
  - For water vapor profile sounding, CLIMCAPS-*Aqua* retrievals produce a smaller MSE than MERRA2 reanalysis by 45% near 600hPa, which highlights the value of additional hyperspectral infrared observations from AIRS to improve the free tropospheric humidity in the reanalysis, while a larger than MERRA2 MSE is found near the surface in the lower troposphere, which corresponds to the decreasing performance of infrared retrieval in these conditions.
  - For temperature profile sounding, CLIMCAPS-*Aqua* MSE is 10% smaller compared to MSE of MERRA2 around 300hPa, which occurs in the tropical regions. This indicates the potential of hyperspectral infrared soundings to further the understanding of thermodynamic conditions associated with tropical convection and large-scale circulation.
  - In terms of sampling bias, CLIMCAPS-*Aqua* generally shows a cold and dry bias in cloudy conditions and a warm and wet bias in less cloudy conditions. The magnitude of the bias and the relationships with clouds depend on the region and surface types. The largest warm and wet bias caused by sampling occurs over snow- and ice-covered land surfaces, while over sea ice CLIMCAPS-*Aqua* generally has a cold and dry bias

regardless of cloud fraction. Over non-frozen ocean, CLIMCAPS-*Aqua* sampling biases peak at a 1-2K warm bias and < 20% wet bias at about 0.5 ECF and a 20% dry bias in the overcast condition. CLIMCAPS-*Aqua* has much reduced number of retrievals when ECF is larger than 0.9 than AIRS V7, therefore, CLIMCAPS-*Aqua* has a larger sampling bias in overcast conditions. The increased yield of CLIMCAPS-*Aqua* in the lower troposphere compared to AIRS V7 results in much reduced sampling biases below 800hPa than AIRS V7.

  - The CLIMCAPS-*Aqua* reported similar spatial distribution of degree of freedom (DOF) for temperature and water vapor with AIRS V7 with larger DOF values in CLIMCAPS-*Aqua*. This difference is partially caused by the different methods to calculate averaging kernels used in the two retrieval systems and the weighting applied in the AIRS V7 for water vapor retrieval discussed in detail by Yue and Lambrigtsen (2019) and Smith and Barnet (2020).

- CLIMCAPS-*Aqua* total column water vapor (TCWV) retrieval compares well with the in-situ observations over land and passive microwave satellite retrievals over ocean. Similar to other AIRS retrieval systems, CLIMCAPS-*Aqua* TCWV has a wet bias in dry conditions (TCWV< 20 mm) and a dry bias in wet conditions (TCWV>60mm) with magnitude within ±20%. All retrieval systems are within a ±5% bias at intermediate TCWV values. CLIMCAPS-*Aqua* has a smaller dry bias associated with deep convective regions such as ITCZ and midlatitude storm tracks than AIRS V7 retrievals. In the low cloud topped subtropical oceanic regions, CLIMCAPS-*Aqua* has a small dry bias compared to a small wet bias in AIRS V7. CLIMCAPS-*Aqua* has higher yields than AIRS V7 in mid- and high-latitude regions. However, CLIMCAPS-*Aqua* has a 20% yield reduction during daytime/nighttime soundings than AIRS V7 over Australia/Central Asia.
- The near surface temperature and humidity retrievals from CLIMCAPS-*Aqua* present good agreements with in-situ observations over both land and ocean, which is likely related to MERRA2 reanalysis used as first guess.
- CLIMCAPS-*Aqua* retrievals are tested in two climatic extreme conditions in this report: polar region and the 2003 summer heat wave event over Europe. In the polar region, CLIMCAPS-*Aqua* seems to produce a larger cold and dry bias in winter conditions than AIRS V7, which is related to sampling differences between CLIMCAPS-*Aqua* and AIRS V7. CLIMCAPS-*Aqua* is able to capture the variability related to extreme events such as heat waves and the retrieval performance is stable across different conditions, although a sudden change in moisture lapse rate in the lower troposphere near the surface is found in the retrieved profiles. The error
- The total column ozone retrievals are evaluated with coincident Ozone Monitoring Instrument (OMI) measurements. The relative bias of CLIMCAPS-*Aqua* retrievals is within ±10% from OMI with a negative bias over Southern Ocean in comparisons using January data. The ozone DOF reported in the V2 CLIMCAPS-*Aqua* is lower than that in AIRS V7, particularly in the winter hemisphere. This is partially due to the additional basis functions coupled with more channels near the peak ozone absorption at 1040 cm^{-1} used in the AIRS V7 ozone retrievals,

which may provide additional resolution (and information) in the cold temperatures of the lower stratosphere near the tropopause.

- CLIMCAPS-*Aqua* cloud retrievals are evaluated by comparing with AIRS Version 6 (V6), which has been validated extensively in previous studies. Results show that CLIMCAPS-*Aqua* global average total ECF agrees with AIRS V6, the regional differences, however, are larger and frequently change sign depending on latitude, the presence of land, ocean, sea or land ice, and cloud regimes.
    - The CLIMCAPS-*Aqua* ECF is smaller and less spatially consistent than AIRS V6 over the low cloud topped subtropical regions and throughout the low latitude surrounding areas of deep tropical convection, with a much colder cloud top temperature (Tcld) in CLIMCAPS-*Aqua*.
    - Tcld of the upper cloud layer in CLIMCAPS-*Aqua* is 5-10K colder than AIRS V6 in the heart of the deep convection in the central Pacific. Many of these clouds are thin cirrus and it could be possible that CLIMCAPS-*Aqua* is better able to assign a correct altitude to these clouds. On the other hand, the lower layer cloud in deep convection is placed at a lower altitude (warmer Tcld) in CLIMCAPS-*Aqua*.

For the performance and quality of the V2 CLIMCAPS-Aqua retrievals in the applications to study various physical processes and investigations on the spatial and temporal variability, this report only includes very limited testing results on these topics, which warrants further analyses.

4. L2 Temperature (T) and Water Vapor (Q) Vertical Profiles

4.1 Summary of Testing Tasks

The V2 CLIMCAPS-*Aqua* L2 T and Q profiles are evaluated by comparisons with pixel-scale collocated European Centre for Medium-Range Weather Forecasts (ECMWF) model analyses and four different in-situ observation datasets including the Integrated Global Radiosonde Archive (IGRA, Durre et al. 2006, Wong et al. 2015), radiosondes during the Marine ARM GPCI Investigation of Clouds (MAGIC) campaign (Kalmus et al. 2015), radiosonde and dropsonde measurements in Arctic and Antarctic regions.

- The comparisons with ECMWF model analysis focus on the vertical, spatial, and cloud-conditioned patterns of retrieval yield and differences between retrieval and EMCWF data in different regions over different surface classes including sampling bias.
- The comparison with IGRA aims at evaluating the T and Q profiles in both the CLIMCAPS-*Aqua* retrieval and initial guess with a focus over the land region since the IGRA stations are mostly over land (Fig. 4.1.1-I).
- The MAGIC campaign (9/2012–10/2013, Fig. 4.1.1-II) included 19 round trips between Los Angeles and Honolulu. The radiosonde data from MAGIC characterized the atmospheric state typical for low cloud topped subtropical region climate regime, where the combined infrared and microwave measurements such as AIRS have shown high retrieval yield and accuracy (Yue et al. 2013, Yue and Lambrigtsen 2019).
- To evaluate the retrieval performance in extreme climatic conditions such as polar region and extreme events, the thermodynamic profiles collected during multiple polar field experiments are used as shown in Table 4.1.1 (Fig. 4.1.1-III) and IGRA sondes over Europe during the 2003 summer heatwave event are used (Fig. 4.1.1-I).

Table 4.1.1: The three Central Arctic field campaigns and the Antarctic Concordiasi field experiments used in CLIMCAPS-*Aqua* polar region retrieval evaluation.

Campaign Name	Geographical Region	Time Period	Sonde Type	Profile/ day	# of Sondes Collocated with AIRS	Reference
ASCOS	77.9°-87.5°N 11.1°W-9.6°E	08/03/2008-09/07/2008	Vaisala RS92	4/Day	75	Tjernström et al., 2014
ACSE	71.4°N-85.2°N 25.7°E-178.1°W	07/06/2014-10/01/2014	Vaisala RS92	4/Day	191	Sotiropoulou et al., 2016
N-ICE2015	79.2°N-83.3°N 3.4°E-29.8°E	01/12/2015-06/22/2015	Vaisala RS92	2/Day	127	Granskog et al., 2016
Concordiasi	Antarctica	09/23/2010-12/08/2010	MIST Dropsonde	12/Day	392	Boylan et al. 2015

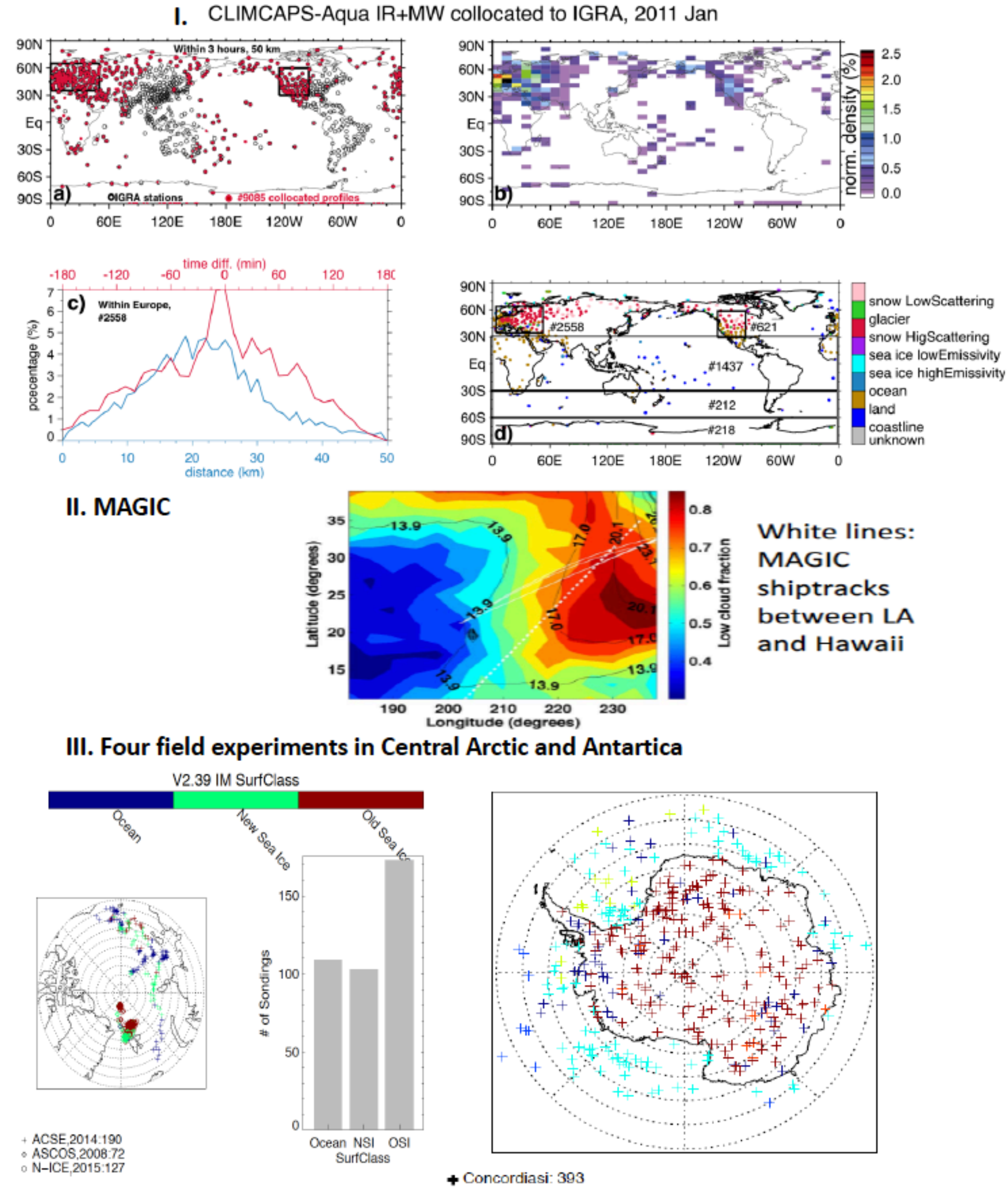


Figure 4.1.1 In-situ observations used to evaluate the CLIMCAPS-*Aqua* L2 T and Q profiles. Panel I shows IGRA radiosondes: (a) The IGRA stations (black circles) and successfully collocated CLIMCAPS-*Aqua* IR+MW profiles (red dots, total 9085) within 3-hour, 50-km searching criteria for January 2011; (b) the density distribution (%) of all collocated profiles; (c) histogram of the distance (lower x-axis, blue) and time differences in minutes (upper x-axis, red) from the collocated 2558 profiles within the Europe region; (d) the collocated profiles color-coded by surface types, with five focused regions in black boxes (Europe [0-50^{o}E,

35-65°N], North America [95-125°W, 30-60°N], Tropics [30°N–S], Southern Ocean [30-60°S], and South Pole [60-90°S]). Data over Europe is used in the heatwave event analysis. Panel II shows MAGIC data used in this analysis with background contours for MODIS mean cloud fraction. Panel III shows the in-situ data for the polar region evaluation: the location of the collocated sondes during the three Arctic field campaigns (indicated by different symbols) specified by the surface class and histograms (left); the location of the Concordiasi dropsondes that are matched with AIRS observations within 50km and 2hr window.

4.2 Summary of Major Findings

4.2.1 Retrieval Yield

Figure 4.2.1.1 summarizes the yield of the V2 retrieval of CLIMCAPS-*Aqua* L2 T and Q profiles. Similar to other infrared and microwave retrieval products, CLIMCAPS-*Aqua* yield is dependent on the surface types and cloud conditions within the field of view (FOV). CLIMCAPS-*Aqua* has a very small number of profiles with QC=1 flags, thus the differences on the yield between QC=0 and QC=0/1 are negligible comparing with AIRS Version 7 (V7) product. CLIMCAPS-*Aqua* yield has much smaller vertical variation below 250hPa, reflecting the "whole profile" quality control of the algorithm. A smaller yield is found in the upper troposphere in partial cloudy conditions compared to V7, but much larger yield between 700hPa and surface. The CLIMCAPS-*Aqua* yield also shows a sudden decrease with effective cloud fraction (ECF) less than 0.1 (0.2) over land and sea ice covered (frozen land) scenes. This point can also be seen clearly from the percentage of "failed" retrievals (QC=2) as given in Figure 4.2.1.2. This is due to the uncertainty in the ECF retrieved by CLIMCAPS-*Aqua* V2 algorithm as shown in Section 8 of this report and Yue et al. (2021). The dependence of yield on various conditions causes the spatial and seasonal variations of the retrieval yield. Fig. 4.2.1.3 shows the zonal mean yield and the percentage of QC=0 cases on January 1, 2011. Comparing with AIRS V7, CLIMCAPS-*Aqua* has a larger yield near the surface and a smaller yield in the free troposphere.

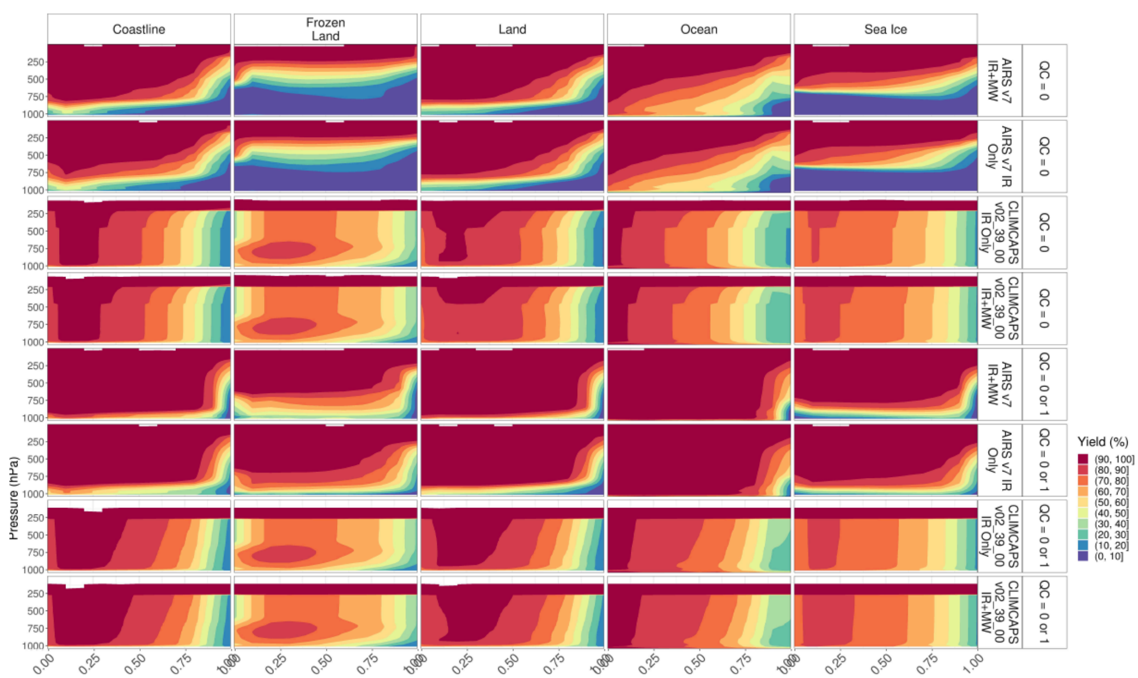


Figure 4.2.1.1. Retrieval yield of various *Aqua*-AIRS L2 temperature and water vapor profile retrieval products varying with surface classes and effective cloud fraction (ECF) values. Results from four

algorithms are shown: AIRS V7 IR+MW, AIRS V7 IR-only, V2 CLIMCAPS-*Aqua* IR+MW, and V2 CLIMCAPS-*Aqua* IR-only. Both QC=0 and QC=0 or 1 flags are shown. The surface classes and ECFs used to bin the data are obtained from AIRS V7 IR+MW algorithm. Results are obtained from one month of retrieval (July, 2011).

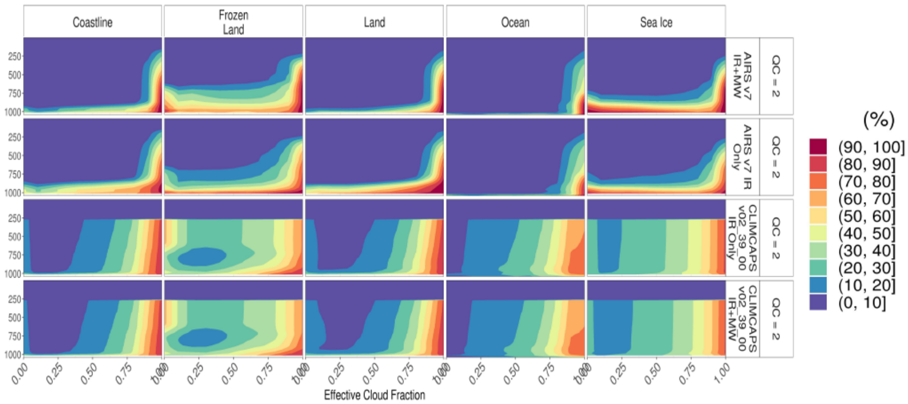


Figure 4.2.1.2. Similar to Figure 4.2.1.2, showing percentage of QC=2 cases instead.

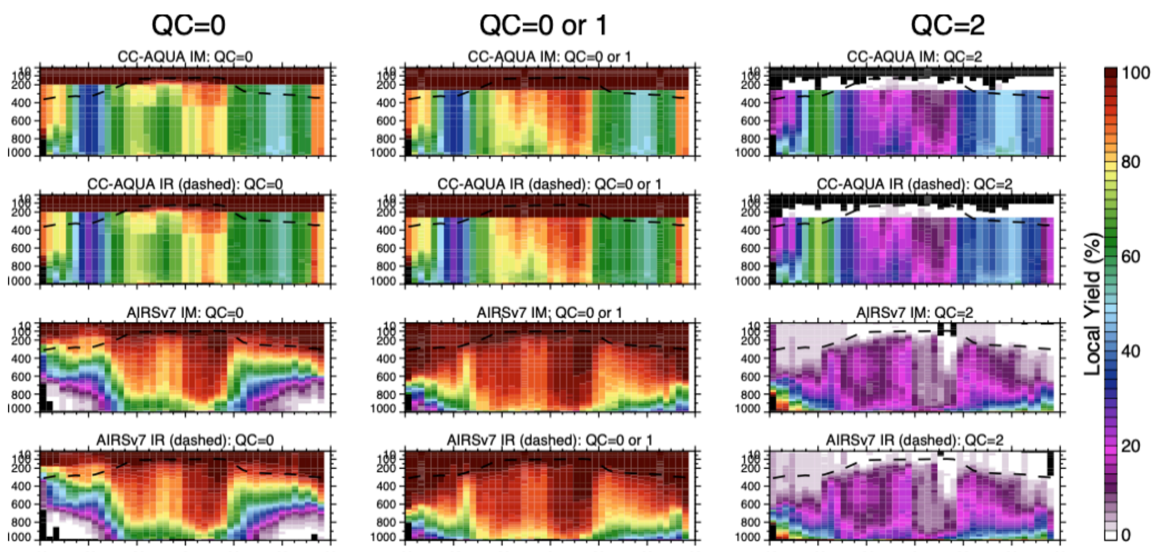


Figure 4.2.1.3. Zonal mean vertical cross sections of the QC frequency of occurrence for CLIMCAPS-*Aqua* IR+MW (CC-*Aqua* IM, top row), CLIMCAPS-*Aqua* IR-only (CC-*Aqua* IR, second row), AIRS V7 IR+MW (AIRSv7 IM, third row), and AIRS V7 IR-only (AIRSv7 IR, bottom row). From left to right, results are shown for quality control flag QC=0 (left column), QC=0 or 1 (middle column), and QC=2 (not for scientific use, right column). The black filling indicating either no data reaching below surface or no data falling into that specific scenario. The black dashed lines mark the lapse-rate tropopause.

4.2.2 Retrieval Biases of L2 Water Vapor Profile

As shown in Figure 4.2.2.1, averaging globally using pixel-scale collocated ECMWF data as reference, the water vapor profiles from V2 CLIMCAPS-*Aqua* algorithms have a similar magnitude of mean bias (within ±10%) with AIRS V7 below 200hPa where AIRS spectra show sensitivity to water vapor. A wet bias is seen below 400hPa, and a dry bias above. Comparing to AIRS V7, CLIMCAPS-*Aqua* produces a much smaller RMSE near the surface and a larger RMSE above 700hPa. Such differences are related to the differences in the first guesses used by the retrievals and larger differences between MERRA2 and ECMWF water vapor profiles found in the upper troposphere. However, the CLIMCAPS-*Aqua* physical retrievals reduce the error of the free troposphere water vapor profile in MERRA2 as the retrieved water vapor profiles produce a much smaller RMSE in the free troposphere than the first guess from MERRA2.

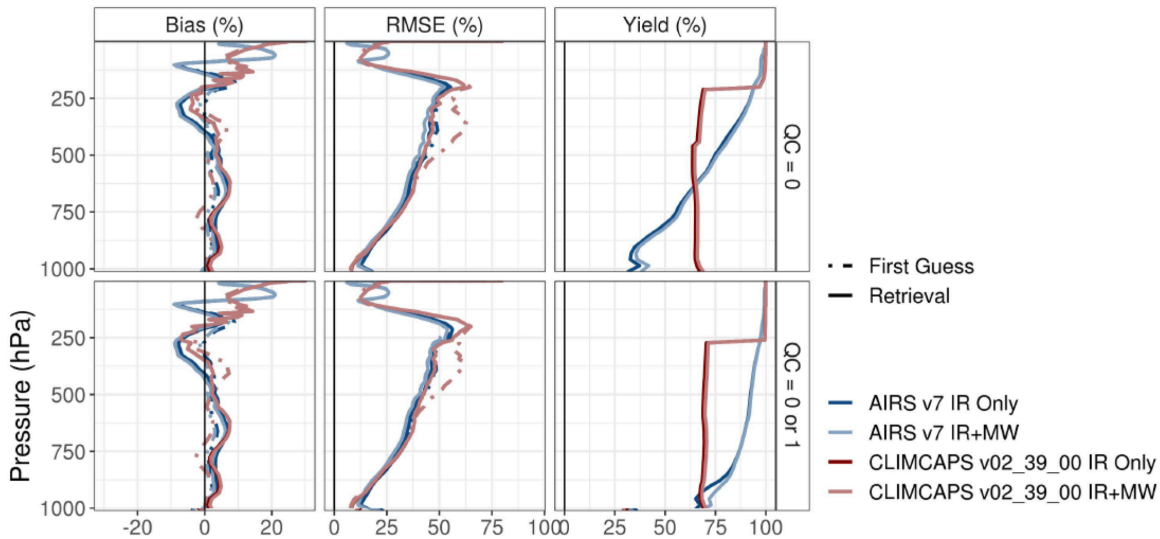


Figure 4.2.2.1. Global mean relative bias (left), RMSE (middle), and yield (right) for water vapor profiles by retrieval system (color), retrieval type (line type), and QC (row) from July 2011. Yield is calculated using the QCs from each retrieval system. Bias and RMSE are calculated from cases filter with CLIMCAPS-*Aqua* IR+MW quality control flags to ensure that statistics are calculated using the same data samples.

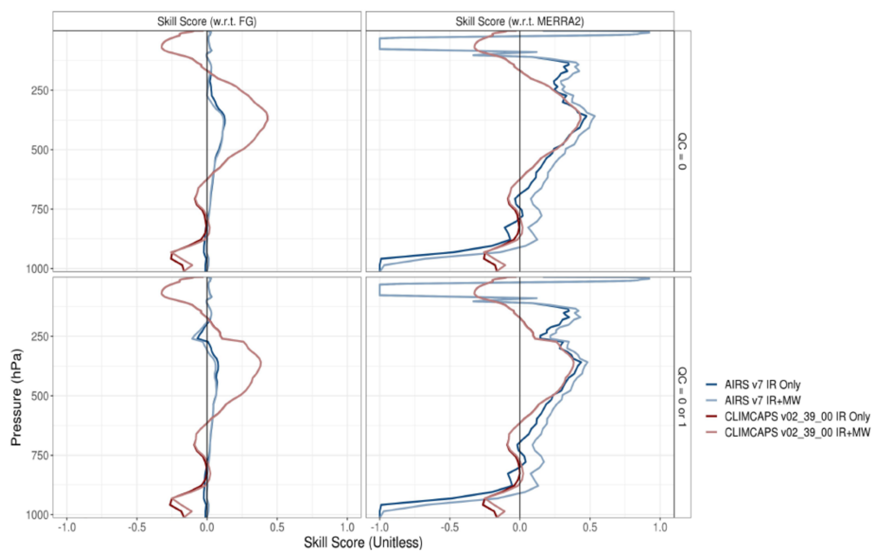


Figure 4.2.2.2. Global mean skill score for Specific Humidity with respect to the first guess (left) and MERRA2 (right) by QC (rows) and retrieval system (color) from July 2011 data. MSEs are calculated from cases filter with CLIMCAPS-*Aqua* IR+MW quality control flags to ensure that statistics are calculated using the same data samples.

The reduction of the magnitude of water vapor MSE comparing to the first guesses and collocated MERRA2 reanalysis data are shown by skill scores calculated using Eqn. (2.7). Skill scores are used to show how the mean square error (MSE) of retrievals differ with the ***target*** data when the same reference dataset is used in the evaluation of MSE. When *target* is first guess of the retrieval, skill scores show the quality of final retrieval (in terms of MSE) with respect to its first guess. Setting *target* to be reanalysis such as MERRA2 shows the quality of retrieval with respect to the reanalysis data. Figure 4.2.2.2 summarizes the global mean MSE comparisons with retrieval first guess and collocated reanalysis for water vapor profiles using ECMWF model analyses as reference data. Averaging globally, CLIMCAPS-*Aqua* retrievals greatly reduce the MSE in the MERRA2 water vapor profiles above 600hPa by nearly 45%, which shows the potential of hyperspectral infrared sounders such as AIRS to further improve the humidity information in the reanalysis.

Comparisons of CLIMCAPS-*Aqua* with IGRA and MAGIC radiosondes (Fig. 4.2.2.3) show results consistent with ECMWF comparisons. The IGRA comparison shows the difference between the CLIMCAPS-*Aqua* IR-only and IR+MW (IM) water vapor profiles is very small with the IR-only algorithm producing slightly larger RMSE near the surface due to the loss of MW information. Only IGRA comparison in the tropics is shown in this report. Wang et al. (2021) includes a more detailed discussion and results in various regions and cloud conditions. The reduced MSE in final retrievals compared with MERRA2 water vapor profiles is apparent in the IGRA study.

Over the subtropical N.E. Pacific region, a dry bias with magnitude less than 20% is observed between 400 hPa and surface in both the final retrieval and the MERRA2 first guess. The maximum bias (1.0 g/kg) and RMSE (2.5 g/kg) occur near the top of the boundary layer around 800hPa where inversions occur frequently in this region.

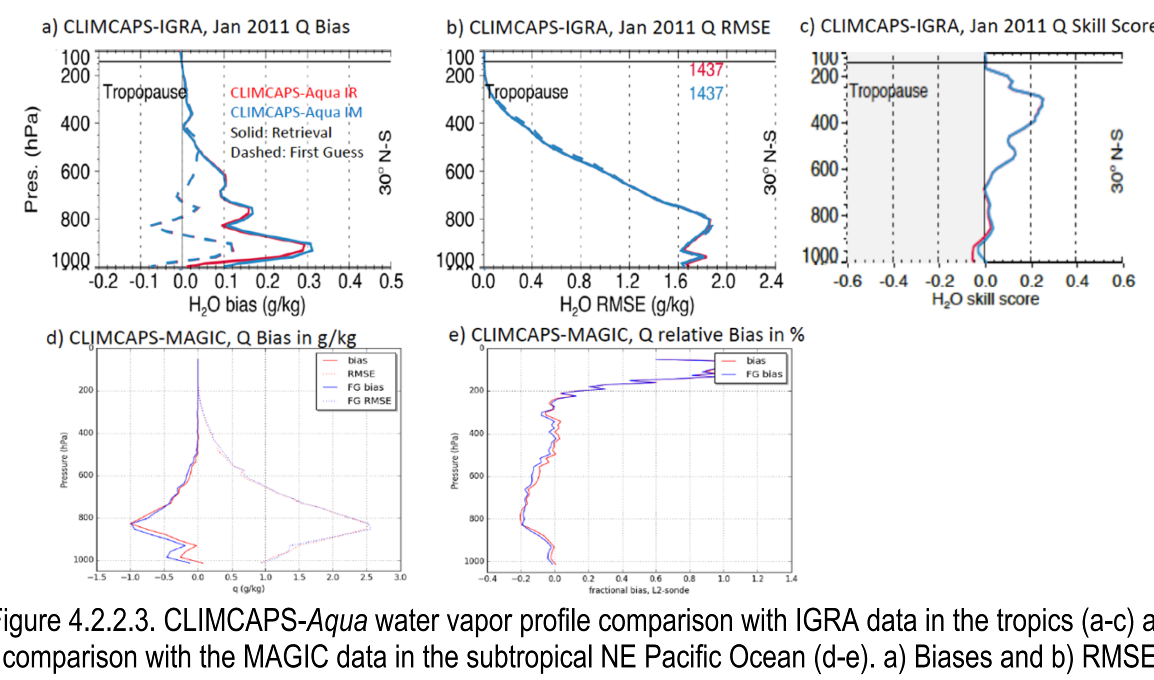


Figure 4.2.2.3. CLIMCAPS-*Aqua* water vapor profile comparison with IGRA data in the tropics (a-c) and comparison with the MAGIC data in the subtropical NE Pacific Ocean (d-e). a) Biases and b) RMSEs comparing to IGRA observations for H_2O (30° N–S) for CLIMCAPS-*Aqua* IR (red) and IR+MW (blue) in January 2011. The solid lines are the CLIMCAPS retrievals minus IGRA, and the dashed lines are the first guess (MERRA2) minus IGRA. C) Skill scores of CLIMCAPS-*Aqua* retrieved water vapor profiles with

respect to MERRA2 using IGRA data as reference. d) Comparisons with MAGIC sondes water vapor bias and RMSE in g/kg for CLIMCAPS-*Aqua* IR+MW retrievals. The relative bias in % is shown in e).

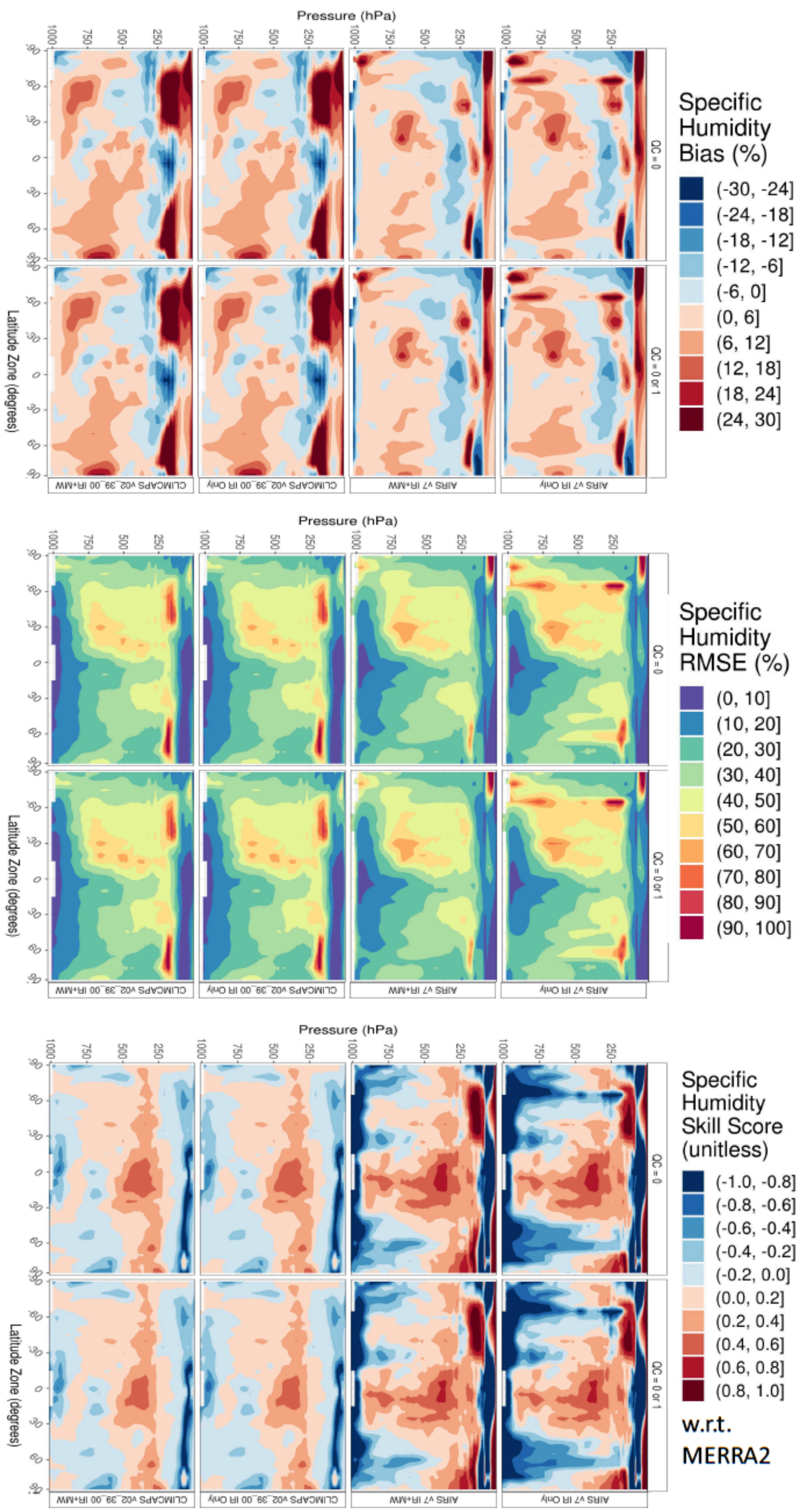


Figure 4.2.2.4: Zonal mean bias (top), RMSE (middle), and skill scores (bottom) for the final retrieval of specific humidity by QC (columns) and retrieval type (rows) using pixel-scale collocated ECMWF data as reference data. July 2011 data is used in the calculation. Positive skill score indicates using ECMWF as reference in the calculation of RMSE, retrievals have smaller RMSEs than collocated MERRA2 reanalysis.

The detailed vertical structure and magnitude of the bias and RMSE vary with season, geographical regions, cloud and surface conditions. However, comparing to AIRS V7, CLIMCAPS-*Aqua* shows smaller variations on the bias and RMSE magnitudes across different conditions, indicating a more stable retrieval performance, which is related to a more stable first guess (MERA2) used in the retrieval. For more details, readers may refer to Wang et al. (2021) for AIRS-IGRA comparisons and Roman et al. (2021) for AIRS-ECMWF comparisons. We summarize the latitudinal variations of water vapor profiles by comparisons with collocated ECMWF model analyses in Figure 4.2.2.4. To ensure same data samples going into the comparison, CLIMCAP-*Aqua* IR+MW quality control flags are applied to filter the different retrievals. Differences between the retrievals above 250hPa are purely due to different first guesses used in the retrieval algorithms since AIRS has limited sensitivity to water vapor above this altitude. CLIMCAPS-*Aqua* has much smaller RMSEs than AIRS V7 near the surface and over the Southern Ocean, and a slightly larger RMSE in the tropical free troposphere. The positive skill scores of water vapor retrievals with respect to MERRA2 reanalysis highlight the potential of additional information from AIRS observations to improve the free tropospheric humidity in the reanalysis, while the negative skill cores near the surface and below the deep convective clouds correspond with the impact of surface and clouds on retrievals in these conditions from the infrared sounder.

4.2.3 Retrieval Biases of L2 Temperature Profile

As shown in Figure 4.2.3.1, averaging globally using pixel-scale collocated ECMWF data as reference, the temperature vertical profiles from V2 CLIMCAPS-*Aqua* algorithms have similar magnitude of mean bias (within ±1K) with AIRS V7, characterized by a cold bias below 700hPa and a warm bias between 300 and 700 hPa. The RMSE magnitudes of CLIMCAPS-*Aqua* are smaller than 2 K throughout the profile and less than 1 K above 600 hPa, which is much smaller than the V7 IR+MW and IR-only retrievals. Similar to the water vapor product, negligible differences are observed between CLIMCAPS-*Aqua* IR-only and IR+MW temperature products.

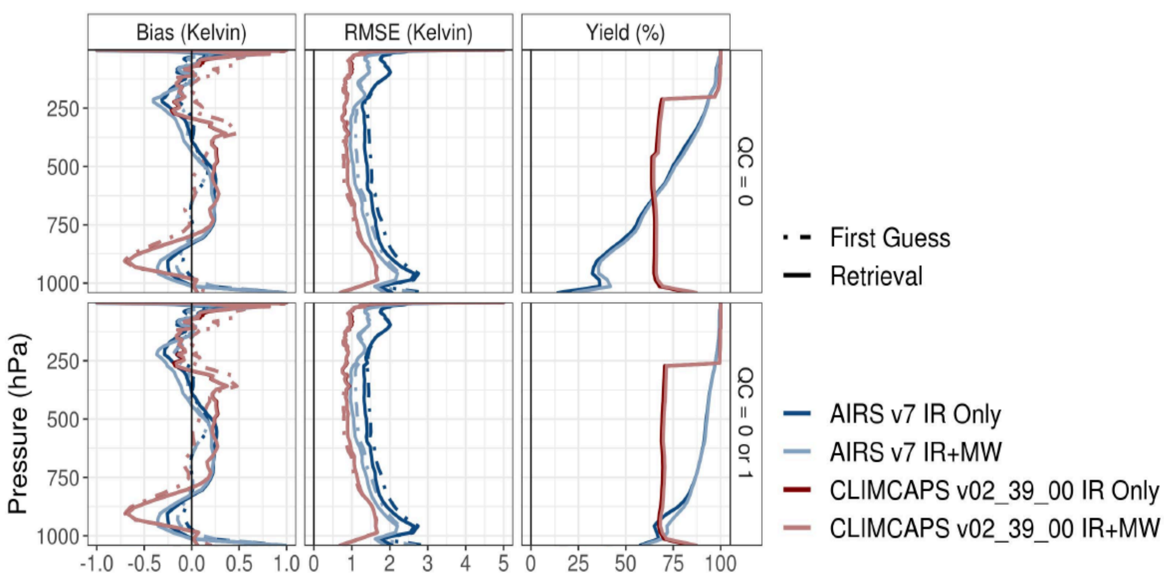


Figure 4.2.3.1. Similar to Fig. 4.2.2.1 but showing results for temperature profile comparison with ECMWF.

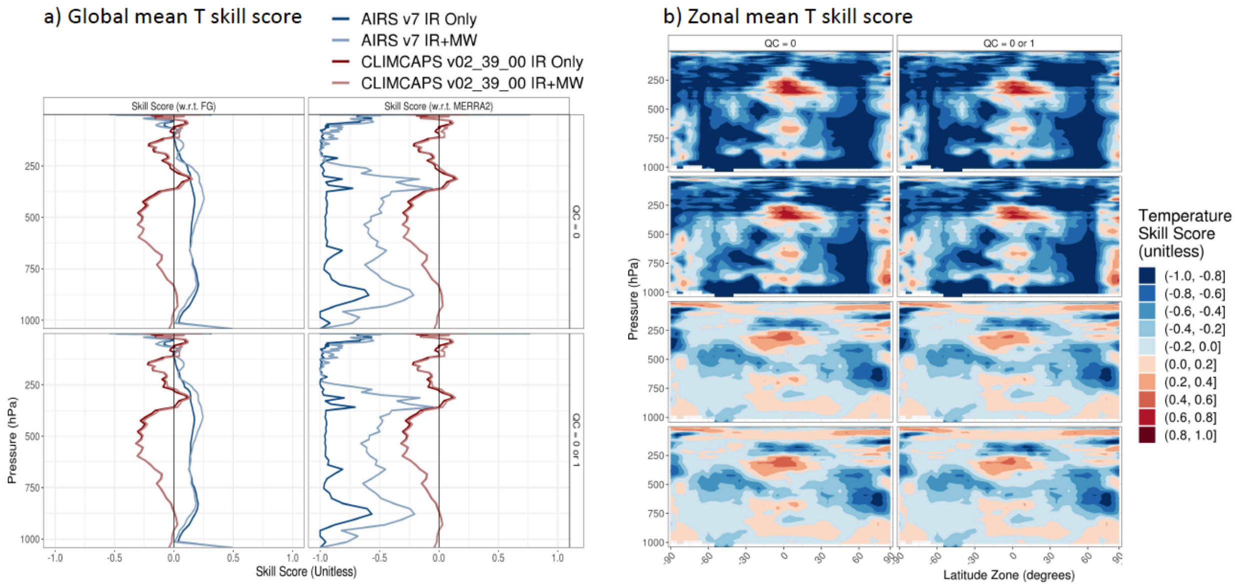
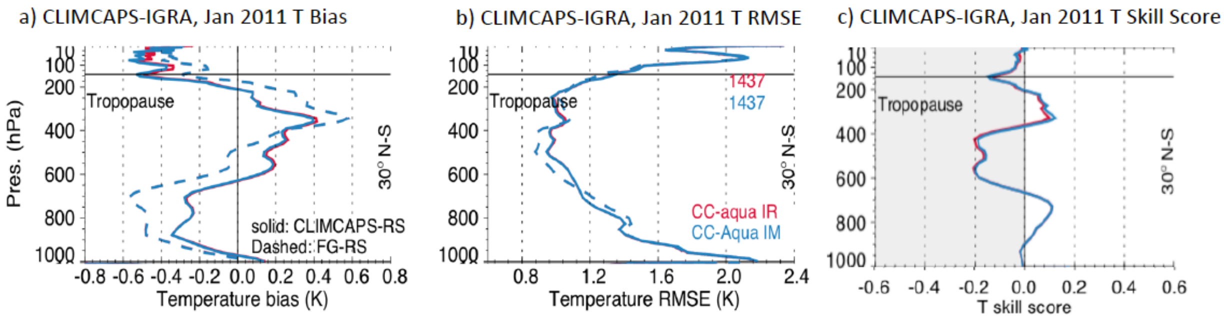


Figure 4.2.3.2. a) Global mean skill score for temperature profile retrievals with respect to the first guess (left) and MERRA2 (right) by QC (rows) and retrieval system (color). B) Zonal mean skill scores for temperature profile with respect to MERRA2 by QC (columns) and retrieval type (rows). The RMSEs are calculated using pixel-scale collocated ECMWF data as reference data from cases filter with CLIMCAPS-*Aqua* IR+MW quality control flags to ensure that statistics are calculated using the same data samples.

Skill scores with respect to each retrieval system's first guess and the MERRA2 reanalysis for temperature are shown in Fig. 4.2.3.2. Different from V7, CLIMCAPS-*Aqua* retrieved temperature profiles show a slightly larger MSE between 400 and 800hPa than the first guess from MERRA2. Indeed, the MSE of CLIMCAPS-*Aqua* retrievals is about 25% larger than MERRA2 in the mid-troposphere, but still much smaller than the MSE of V7, which uses neural net based first guess. CLIMCAPS-*Aqua* retrievals also show a 10% reduction of MSE compared to MERRA2 around 300hPa, which corresponds with positive skill scores in the tropical regions at this altitude (see Fig. 4.2.3.2b and Roman et al. 2021). This indicates the potentials of hyperspectral infrared soundings to further the understanding of thermodynamic conditions associated with tropical convection and large-scale circulation.

The ECMWF comparison results for CLIMCAPS-*Aqua* retrievals are further validated using the in-situ observations from radiosondes from IGRA and the MAGIC campaign (Fig. 4.2.3.3). Since IGRA sondes are mostly over land, the RMSE calculated against IGRA near the surface is larger than the MAGIC and global mean ECMWF comparisons, approaching 2.3K in the tropics. The magnitudes of the CLIMCAPS-*Aqua* lower tropospheric temperature profile biases and RMSEs vary with season and geographical region as discussed in detail by Wang et al. (2021). The MAGIC sonde comparison shows the largest RMSE for CLIMCAPS-*Aqua* temperature retrievals is found near the top of the boundary layer where inversions frequently occur, reaching 1.7 K around 800 hPa.



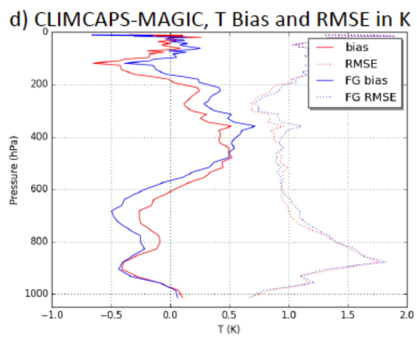


Figure 4.2.3.3. CLIMCAPS-*Aqua* temperature profile comparison with IGRA data in the tropics (a-c) and comparison with the MAGIC data in the subtropical NE Pacific Ocean (d). a) Biases and b) RMSEs comparing to IGRA observations for CLIMCAPS-*Aqua* IR (red) and IR+MW (blue) in January 2011. The solid lines are the CLIMCAPS retrievals minus IGRA, and the dashed lines are the first guess (MERRA2) minus IGRA. C) Skill scores of CLIMCAPS-*Aqua* retrieved T profiles with respect to MERRA2 using IGRA data as reference. d) Comparisons with MAGIC sondes water vapor bias and RMSE in K.

4.2.4 Sampling Biases of the CLIMCAPS-*Aqua* T and Q retrievals

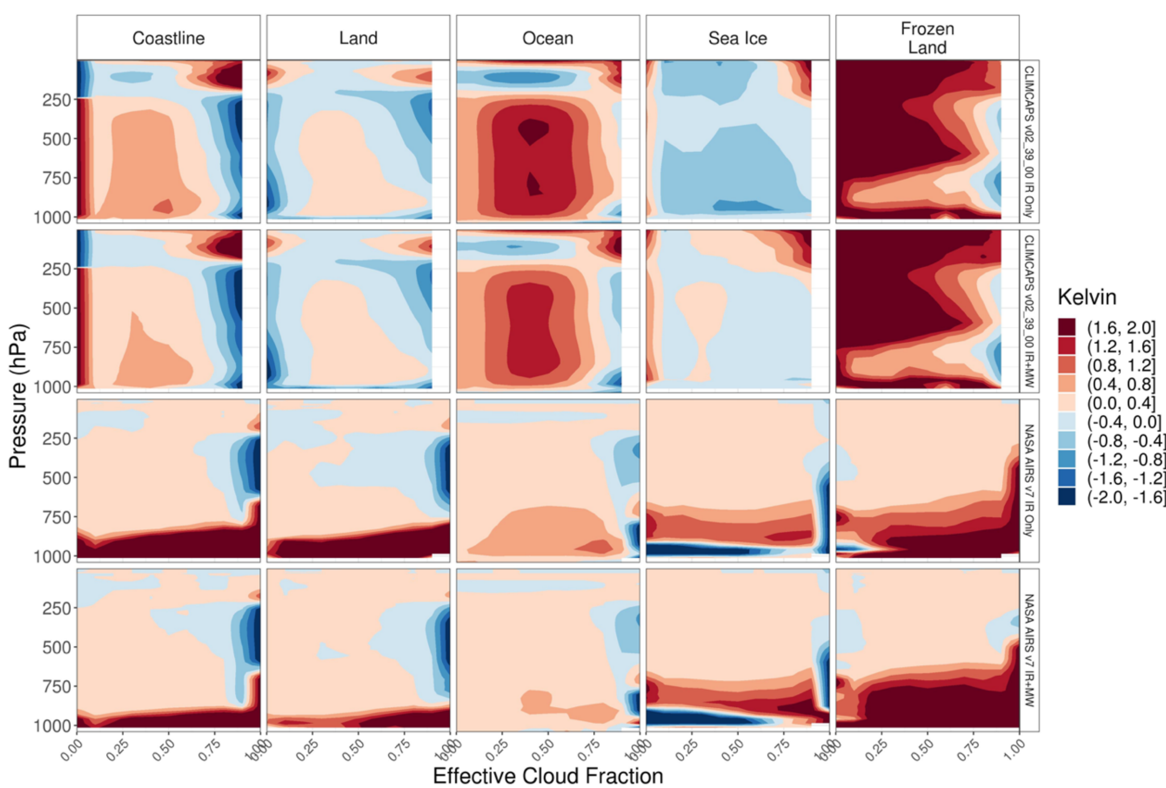


Figure 4.2.4.1. Sampling bias for temperature profiles by effective cloud fraction (x-axis), pressure (y-axis), surface type (columns), and retrieval system (rows). Results from four algorithms are shown: AIRS V7

IR+MW, AIRS V7 IR-only, V2 CLIMCAPS-*Aqua* IR+MW, and V2 CLIMCAPS-*Aqua* IR-only. The surface classes and ECFs used to bin the data are obtained from AIRS V7 IR+MW algorithm. Results are obtained from one month of retrieval (July, 2011).

Previous studies have demonstrated that the cloud-state-dependent sampling bias is the dominant term of total bias in the retrievals obtained from AIRS (Yue et al. 2013; Tian et al. 2020) and the magnitude of the sampling bias also depends on the surface types of the pixel (Yue and Lambrigtsen 2019). Thus, the CLIMCAPS-*Aqua* sampling bias is estimated using collocated ECMWF model analyses following the method in Yue et al. (2013) for temperature and water vapor profiles, respectively. In this study, 7 days of data during July 2011 are used. As shown in Figs. 4.2.4.1 and 4.2.4.2, CLIMCAPS-*Aqua* retrievals generally show a cold and dry bias in cloudy condition and a warm and wet bias in less cloudy conditions caused by sampling. The magnitude of the bias and the relationships with clouds depend on the surface classes. For example, the largest warm and wet bias occurs over snow- and ice-covered land surfaces, while over sea ice CLIMCAPS-*Aqua* generally has a cold and dry bias regardless of cloud fraction. Over non-frozen ocean, CLIMCAPS-*Aqua* shows a 1-2 K warm bias and <20% wet bias at about 0.5 ECF and a 20% dry bias in the overcast condition due to sampling. CLIMCAPS-*Aqua* has a much reduced number of retrievals when ECF is larger than 0.9. AIRS V7 has a smaller sampling bias in cloudy cases, but much larger below 800hPa due to decreasing yield (See Sec 4.2.1).

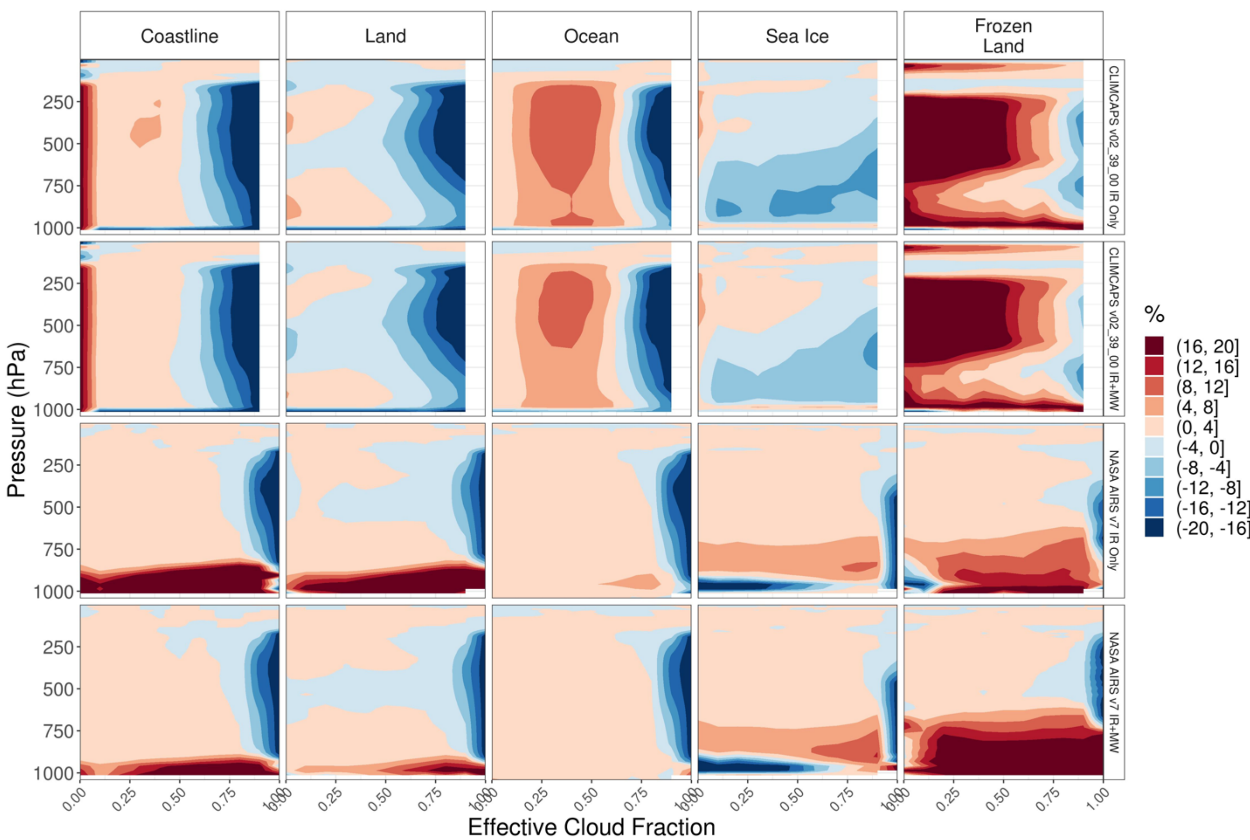


Figure 4.2.4.2. Sampling bias for water vapor profiles by effective cloud fraction (x-axis), pressure (y-axis), surface type (columns), and retrieval system (rows). Results from four algorithms are shown: AIRS V7 IR+MW, AIRS V7 IR-only, V2 CLIMCAPS-*Aqua* IR+MW, and V2 CLIMCAPS-*Aqua* IR-only. The surface

classes and ECFs used to bin the data are obtained from AIRS V7 IR+MW algorithm. Results are obtained from one month of retrieval (July, 2011).

4.2.5 T and Q Profile Retrievals in Polar Region

It is challenging to retrieve the vertical structures of temperature and water vapor in the polar region using measurements from AIRS due to the complex surface radiative property and lack of thermal contrast. Validation studies of AIRS L2 temperature and water vapor profiles in the high latitude region have been scarce especially in the central Arctic and Antarctica, where in-situ measurements from ground-based stations are highly limited. In-situ radiosonde and dropsonde measurements from four different field experiments have been collected and collocated with the AIRS L2 retrievals to evaluate the performance of CLIMCAPS-*Aqua* products in these regions (see Table 4.1.1 and Fig. 4.1.1 III for details).

Fig. 4.2.5.1 demonstrates the retrieval yield in the polar region by various AIRS retrieval systems. The IR+MW algorithms produce higher yields than IR-only since microwave information is used to determine the surface types that are particularly important in the high latitude retrieval. This yield difference is larger for CLIMCAPS-*Aqua* system than for AIRS V7. Old Sea Ice type in the Arctic comparison (snow/ice covered land for Antarctica comparison, not shown) has a higher yield than new sea ice surfaces, which warrants further investigation. CLIMCAPS-*Aqua* retrieval yield does not seem to depend on the cloud state determined by the algorithm over snow- and ice-covered surfaces, which is different from AIRS V7. CLIMCAPS-*Aqua* and AIRS V7 produce similar surface types, therefore, the differences are not caused by different surface classifications in the algorithms.

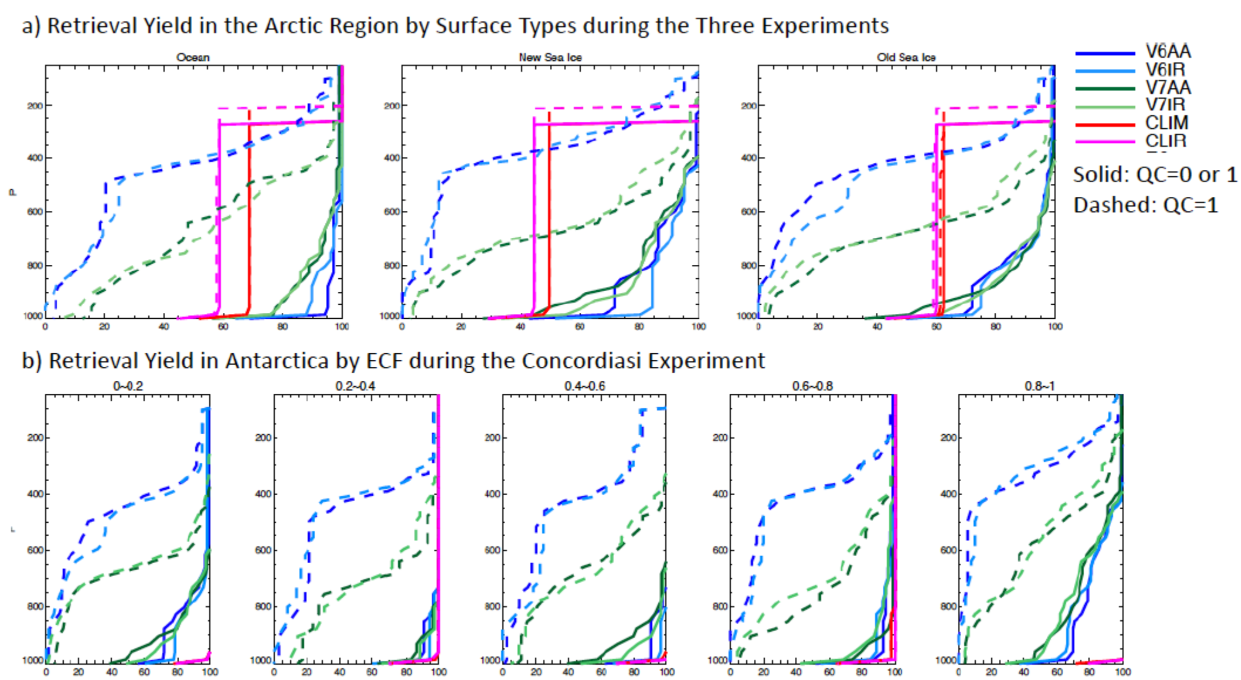


Figure 4.2.5.1 Retrieval yield during the three Arctic Field experiments by surface classifications (a) and by effective cloud fractions (ECF) from AIRS during the Concordiasi experiment in the Antarctica (b). AIRS V7 IR+MW (V7AA), AIRS V7 IR-only (V7IR), V2 CLIMCAPS-*Aqua* IR+MW (CLIM), and V2 CLIMCAPS-*Aqua* IR-only (CLIR) are shown by different colors.

Figs. 4.2.5.2 and 4.2.5.3 show the mean biases in the Arctic region in different seasons for temperature and water vapor, respectively. To demonstrate the differences caused by the quality control indicators in different retrieval systems, results obtained with two approaches are shown. The first one uses the QC flags from each individual product, which evaluates data as a regular data user (Panel a). The second one applies the V7 QC flag to all retrieval products to ensure same samples used in the comparison to evaluate the retrievals (Panel b). As expected, larger magnitudes of biases in these challenging conditions are seen compared to the global mean results and evaluations at more temperate regions shown in previous sections. However, the magnitudes of the biases from different retrieval systems are similar, all showing a cold and dry bias in winter. CLIMCAPS-*Aqua* produces a larger bias than AIRS V7 when using its own QC, which is purely due to the sampling biases caused by the quality control indicators in the two retrieval systems. Therefore, in Fig. 4.2.5.3 the mean biases in the Antarctica region are calculated using the V7 QC flags and binned by the V7 ECF. The performance of different AIRS retrievals in this region is similar to each other with CLIMCAPS-*Aqua* generally produces a smaller bias in water vapor profiles. Below 800hPa, CLIMCAPS-*Aqua* has a larger cold bias than AIRS V7 and collocated ECMWF model analyses, which is due to its first guess, MERRA2, being colder than dropsonde measurements in this region. For a more detailed discussion on high latitude AIRS retrieval validation, readers may refer to Yue et al. (2021).

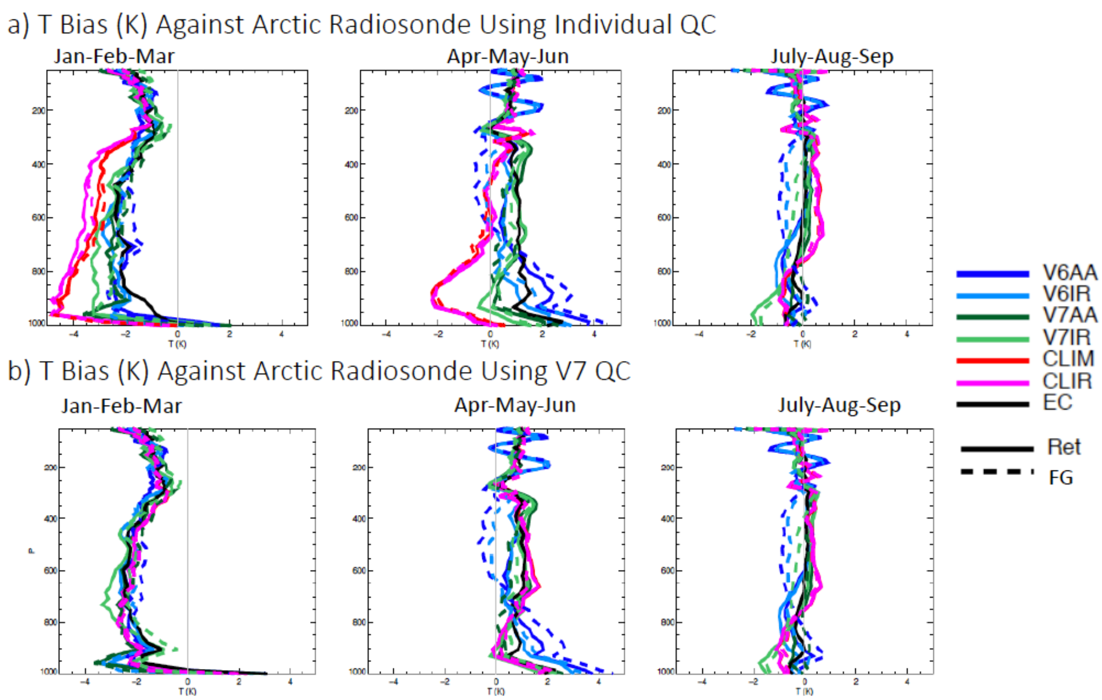


Figure 4.2.5.2 Mean bias in K for temperature (T) profiles of various AIRS retrieval products and ECMWF model analyses as indicated by different colors in different seasons evaluated using the three Arctic Field experiments. The sold and dashed lines correspond with biases in the final retrieval and the first guess (FG) used in different systems: AIRS V7 SCCNN, and MERRA2 for CLIMCAPS-*Aqua*. a) results using QC flags from each individual algorithm, b) results using QC flags from AIRS V7 IR+MW retrievals. AIRS V7 IR+MW ECF and QC flags are used in the calculation. AIRS V7 and V6 IR+MW (V7AA and V6AA), AIRS V7 and V6 IR-only (V7IR and V6IR), V2 CLIMCAPS-*Aqua* IR+MW (CLIM), and V2 CLIMCAPS-*Aqua* IR-only (CLIR) are shown by different colors. EC represents ECMWF model analyses collocated to AIRS.

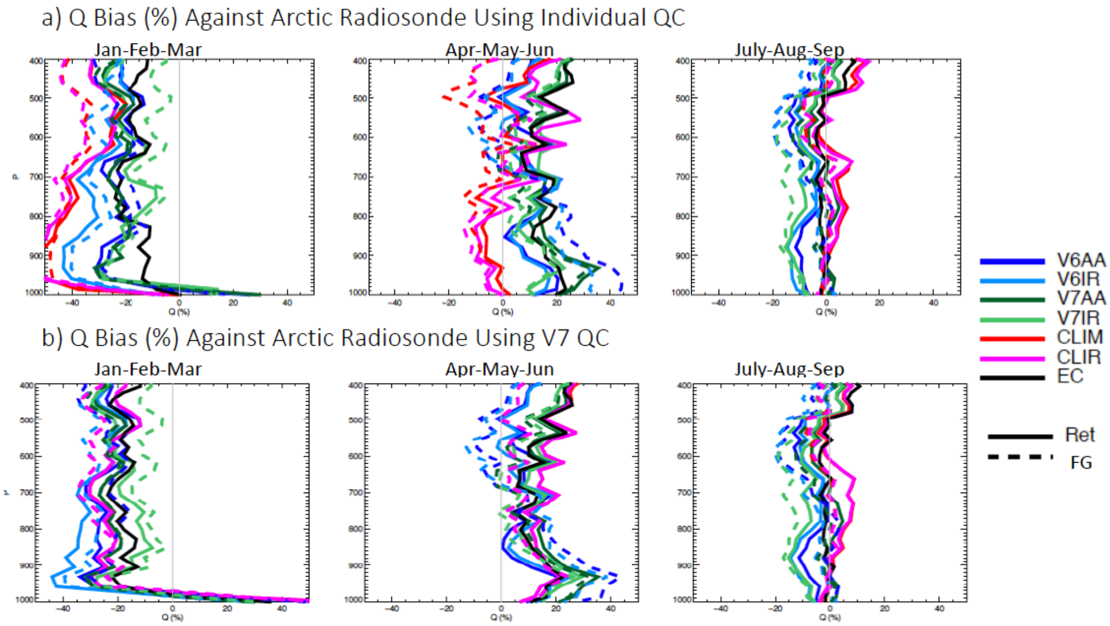


Figure 4.2.5.3 Similar to Fig. 4.2.5.2 but showing relative mean bias of water vapor profiles.

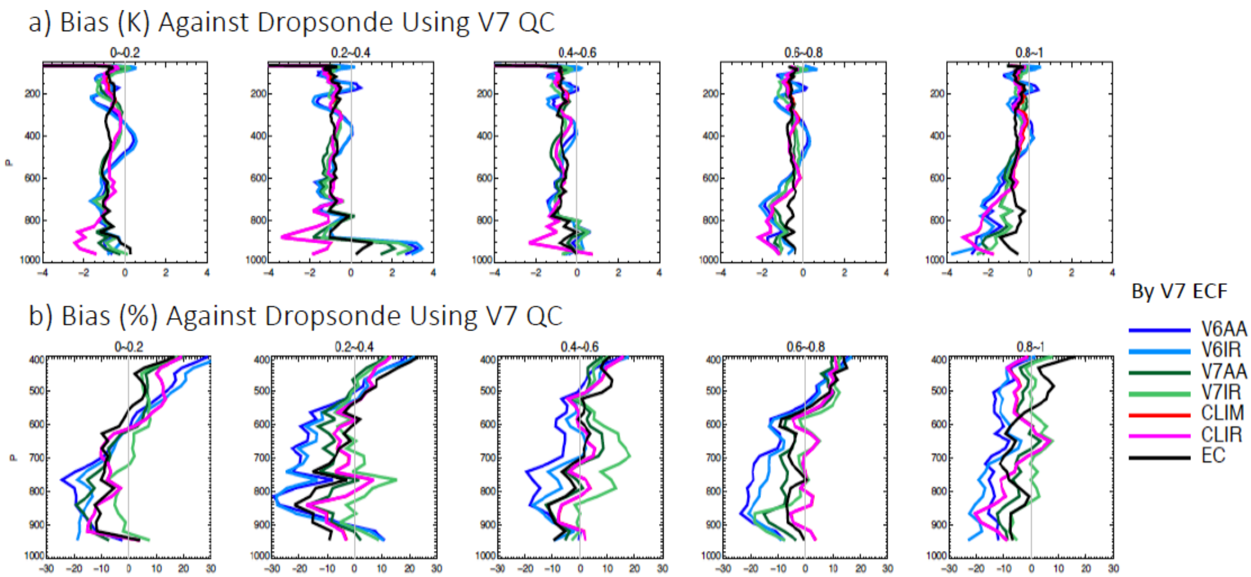


Figure 4.2.5.4 Mean bias for temperature (T in K, a) and water vapor (relative bias in %, b) profiles of various AIRS retrieval products and ECMWF model analyses as indicated by different colors in different effective cloud fraction (ECF) bins calculated against the Concordiasi Experiment in the Antartica. The sold and dashed lines correspond with biases in the final retrieval and the first guess (FG) used in different systems: AIRS V7 SCCNN, and MERRA2 for CLIMCAPS-*Aqua*. AIRS V7 IR+MW ECF and QC flags are used in the calculation. AIRS V7 and V6 IR+MW (V7AA and V6AA), AIRS V7 and V6 IR-only (V7IR and V6IR), V2 CLIMCAPS-*Aqua* IR+MW (CLIM), and V2 CLIMCAPS-*Aqua* IR-only (CLIR) are shown by different colors. EC represents ECMWF model analyses collocated to AIRS.

4.2.6 T and Q Profile Retrievals in the 2003 Europe Summer Heat Wave Event

In late July of 2003 a heat wave occurred over the European continent while July of 2011 is closer to the climatological normal. Comparing the T and Q soundings of these two years provides an opportunity for testing retrievals against extreme hot and dry conditions. Biases and RMSEs of T and Q retrievals from CLIMCAPS-*Aqua* V2 with and without the AMSU microwave channels are evaluated against collocated radiosonde measurements from the IGRA radiosondes for different ranges of near-surface temperatures measured by the radiosondes. Retrievals from AIRS V7 (with and without the AMSU microwave channels) are also shown alongside those from CLIMCAPS-*Aqua* so that performance from the two different retrieval systems can be referenced.

It is found that T profiles in the lower troposphere are close to radiosonde measurements (within ~1 K in most cases) for both V7 and CLMCAPS-*Aqua* retrievals (with or without AMSU channels), implying that both retrieval systems can capture the temperature variability in normal and heatwave conditions. Q profiles in the lower troposphere have wet biases in both retrievals. For AIRS V7 the biases are larger in the year with heatwave compared to the normal year. For CLMCAP-*Aqua* the wet biases are not dependent on whether it is a heatwave or normal year, but suddenly drop to close to zero near the surface, creating a sudden change in moisture lapse rate around 850-900 hPa. In the lower troposphere, AIRS V7 shows a larger RMSE in the heat wave event than the normal year, whereas CLIMCAPS-*Aqua* performs similarly in the two years. Moreover, the error estimates reported in the CLMCAP-*Aqua* L2 product are more comparable to the RMSEs against radiosonde measurements than the error estimates reported in AIRS V7.

Figure 4.2.6.1 shows the T profiles of different retrievals binned at 4 ranges of surface temperatures as well as the radiosonde T profiles. The sample sizes in 2003 July are largest for the surface temperature range of 295-300 K, while those in 2011 July are largest in a wider surface temperature ranges of 290-300 K, implying the 2003 heatwave shifted the temperature histogram to warmer ranges. Both CLIMCAP-*Aqua* and AIRS V7 can capture the radiosonde T profiles and the biases are in general within ±1 K, regardless whether it is in the year of heatwave or not. Also noted is that the differences between CLIMCAPS-*Aqua* IR-only and IR+MW retrievals are very small (so that only red lines are seen).

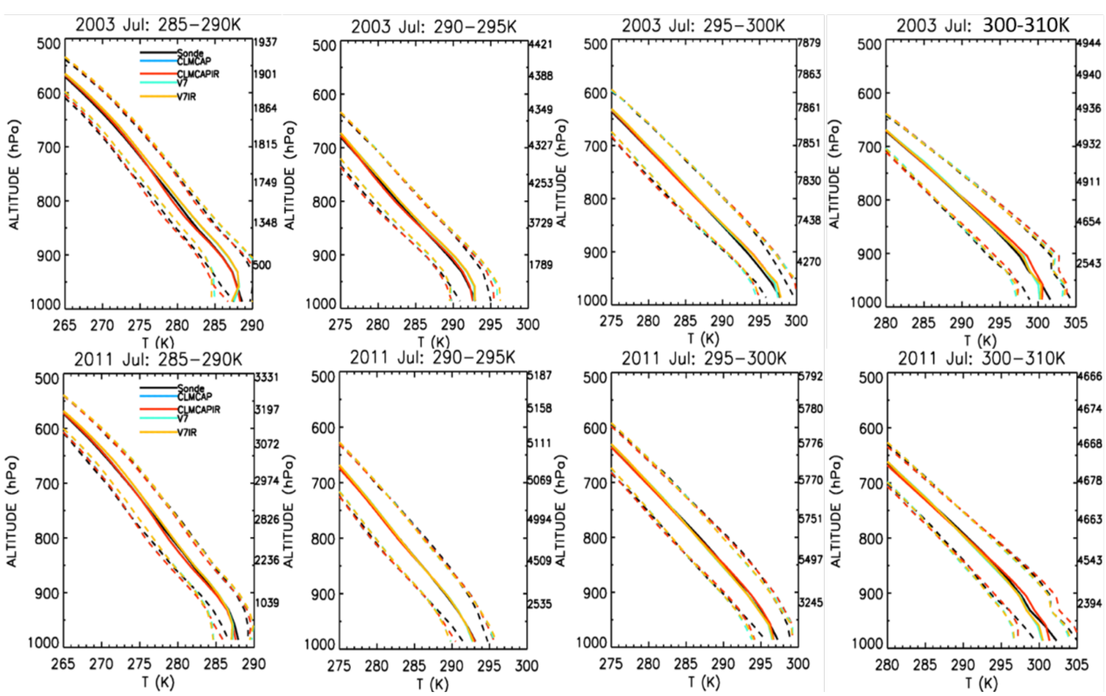


Figure 4.2.6.1 T profiles (in K) averaged over the European region for (top) 2003 July and (bottom) 2011 July. The solid lines are T profiles of the radiosonde measurements (black), CLIMCAP-*Aqua* retrievals with

and without AMSU microwave channels (CLMCAP, blue, and CLMCAPIR, red, respectively), and AIRS V7 retrievals with and without the microwave channels (V7, turquois, and V7IR, yellow, respectively). The dashed lines are the ±1 standard deviations. The profiles are binned for surface T ranges of 285-290 K (1st column), 290-295 K (2nd column), 295-300 K (3rd column), and 300-310 K (4th column). The numbers on the right axes show the sample sizes. All profiles are selected by the same quality control flags (of V7IR) of 0 and 1.

While the CLIMCAPS-*Aqua* retrievals of T can capture the variability in both normal and heatwave conditions, Q retrievals have larger positive biases (moister than the radiosonde measurements) in the lower troposphere during the heatwave year as shown in Fig. 4.2.6.2. This is because the retrievals cannot capture the anomalous dry conditions (Fig. 4.2.6.2 top), particularly for higher surface temperature ranges. There are discrepancies between different retrievals. While AIRS V7 (IR-only and IR+MW) retrievals have moist biases throughout the lower troposphere in 2003 July, CLIMCAPS-*Aqua* shows a reduced moist bias below 750 hPa, and the Q retrievals are much closer to the radiosonde measurements below 900 hPa. However, this vertical bias structure change also results in sudden changes in moisture lapse rates around 850 hPa in the CLIMCAPS-*Aqua* Q retrievals. The retrieval biases of Q profiles (all colored lines in Fig. 4.2.6.2 subtract the black lines) are shown in Fig. 4.2.6.3. Although CLIMCAPS-*Aqua* Q relative biases increase with surface temperature, the biases are less dependent on whether it is in the heatwave year or normal year. On the other hand, both AIRS V7 retrievals have larger moist biases in the heatwave year compared to the normal year. The sudden changes in moisture lapse rate around 850 hPa in the CLIMCAPS-*Aqua* retrievals (seen in Fig. 4.2.6.2) correspond to the sudden drop in moist biases below 800 hPa (seen in Fig. 4.2.6.3). This feature is common in the CLIMCAPS-*Aqua* Q retrievals for all surface temperature ranges and for both heatwave and normal years.

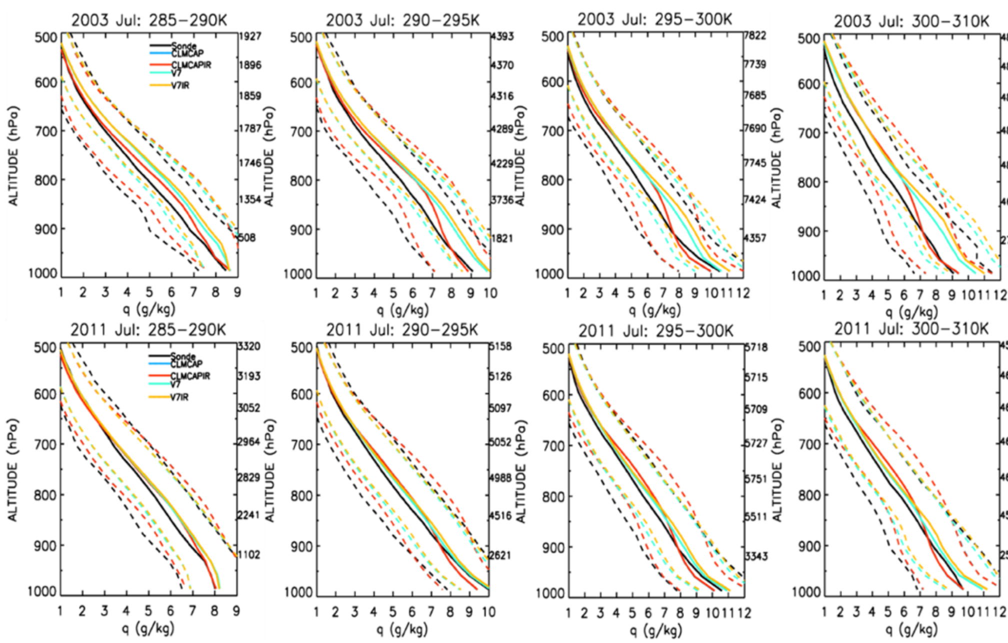


Figure 4.2.6.2 Similar to Fig. 4.2.6.1, but for q profiles (in g/kg).

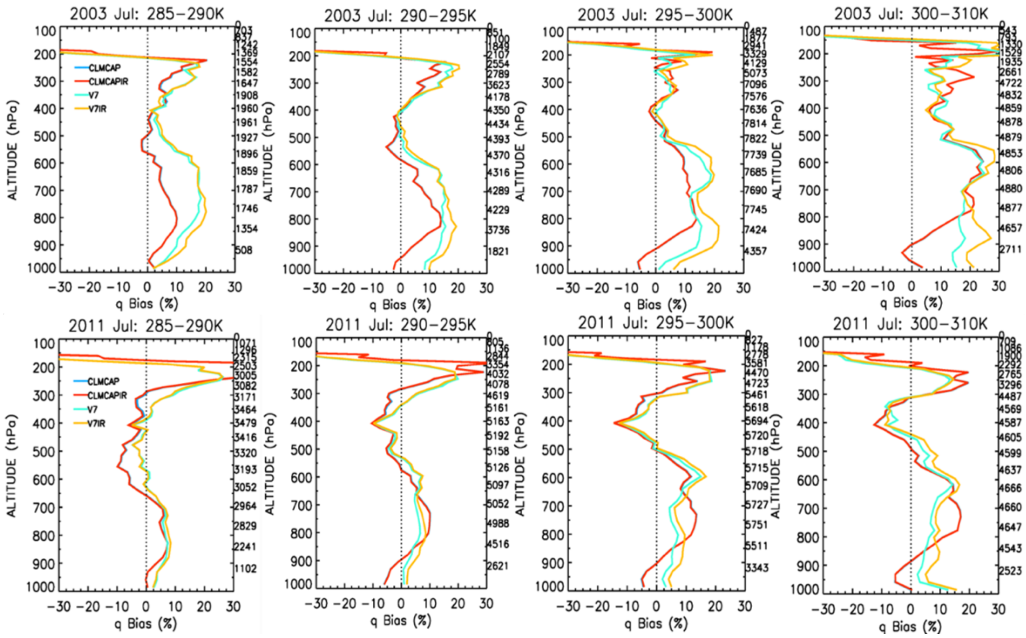


Figure 4.2.6.3 Relative biases (in %) of CLIMCAPS-*Aqua* (red and blue lines for with and without AMSU microwave channels, respectively) and AIRS V7 (turquois and yellow lines for with and without AMSU microwave channels) q retrievals in 2003 July (top) and 2011 July (bottom). The numbers on the right axes are sample sizes (selected by V7IR quality flags of 0 or 1). Both CLIMCAPS-*Aqua* retrievals are very close so that the red and blue lines are overlapped. Also note that the vertical coordinates are from 100 to 1000 hPa.

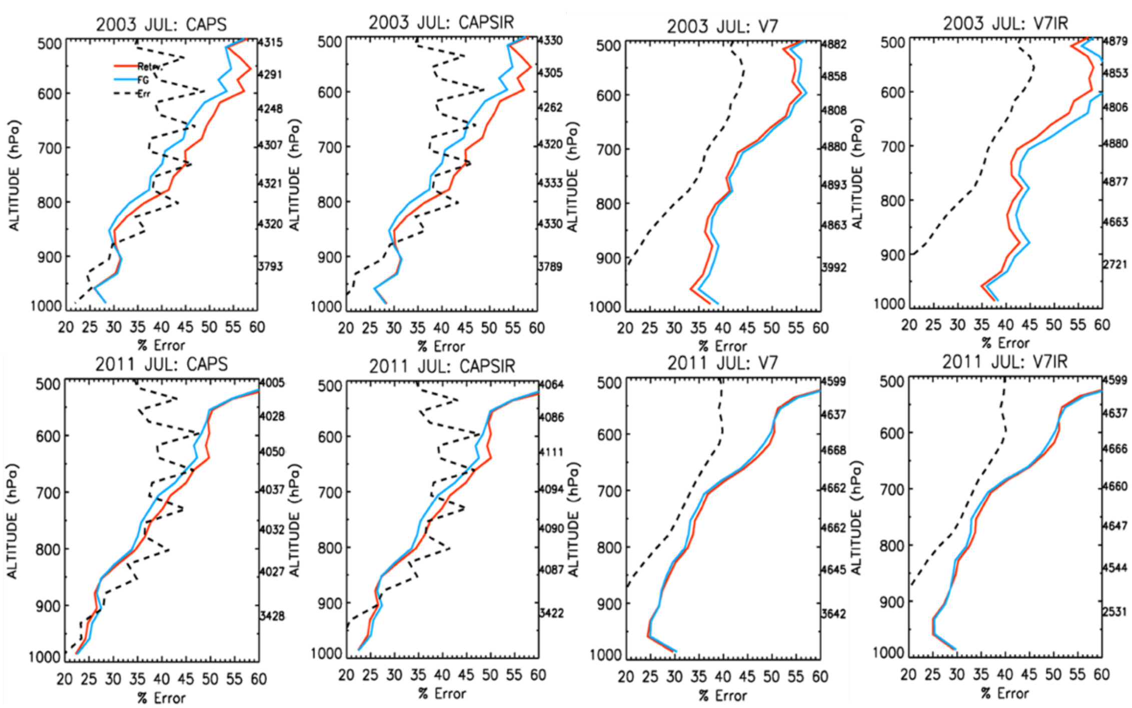


Figure 4.2.6.4 Relative root-mean-squared errors (in %) against radiosonde measurements for the fist guesses (FG, blue lines) and retrievals (Retrv., red lines) of q for (top) 2003 July and (bottom) 2011 July in the surface temperature range of 300-310 K. The first column (CAPS) is for the CLIMCAPS-*Aqua* with the AMSU channels, the second column (CAPSIR) is for the CLIMCAPS-*Aqua* without the AMSU channels, the

third column (V7) is the AIRS V7 with the AMSU channels, and the fourth column (V7IR) is the AIRS V7 without the AMSU channels. The black dashed line (Err) in each panel is the error estimates reported in the corresponding data product files. The numbers on the right axes are sample sizes for individual retrievals.

To understand the behaviors of different Q retrievals in the heat wave year, the Q RMSEs for different retrievals are shown in Fig. 4.2.6.4 as well as the RMSEs of their corresponding first guesses against the radiosonde measurements for the hottest surface temperature range (300-310 K). In a normal year (2011 July), both CLIMCAPS-*Aqua* retrievals have slightly larger RMSEs than their first guesses above 800 hPa, but stick very closely to the first guesses below 800 hPa, while both AIRS V7 IR-only and IR+MW retrievals are very close to the first guesses. In the heat wave year (2003 July), both AIRS V7 retrievals slightly reduce the RMSEs from their first guesses, and the CLIMCAPS-*Aqua* retrievals show larger RMSEs than their first guesses above 850 hPa and stick to the first guesses below 850 hPa. Since CLIMCAPS-*Aqua* utilizes MERRA2 reanalysis Q as first guess, which has assimilated the radiosonde measurements and shows high quality T and Q profiles in regions with dense surface observations, further improvement from reanalysis in these regions is challenging. On the other hand, AIRS V7 utilizes neural network as the first guesses, which have larger RMSEs from the radiosonde measurements especially during the heat wave year due to the limited variability associated with extreme events such as heat wave captured during the neural network training (Yue and Lambrigtsen 2019); therefore, a much larger retrieval improvement from the first guesses is seen when comparing with IGRA. Given differences in the algorithms and first guesses used, CLIMCAPS-*Aqua* have comparable RMSEs to AIRS V7 above ~800 hPa, because of large sensitivity of AIRS spectra in the free troposphere, whereas the smaller RMSEs than AIRS V7 near the surface are largely due to the first guess differences.

Figure 4.2.6.4 compares the error estimate reported in the L2 product for T and Q profiles with the RMSE calculated using the radiosonde observations. CLIMCAPS-*Aqua* estimates are more comparable to the RMSEs against radiosondes below 650 hPa compared to AIRS V7, which significantly underestimate the errors.

5. L2 Total Column Ozone

5.1 Summary of Testing Tasks

The L2 total column zone from V2 CLIMCAPS-*Aqua* retrievals are compared with Ozone Monitoring Instrument (OMI) data. Two days of data are used: January 1, 2011, and July 1, 2011. OMI is on the Aura platform (about 16 minutes ahead of AIRS on *Aqua*), and the Version 3 of the "Differential Optical Absorption Spectroscopy" (DOAS) Level 2 product (OMDOAO3) was used (See OMI Team, 2012, and Veefkind et al., 2006). OMI has a 13x24 km footprint at nadir while the AIRS L2 nadir footprint is approximately a circle of 45 km diameter. Comparisons were one-to-one, with the AIRS observation of quality flag 0 (best) or 1 (good) geographically closest to an OMI footprint selected, but not more than 60 km away. As OMI relies on backscattered UV radiation, only sunlit measurements of AIRS could be used. Quality-passed matchups were not usually available over ice. AIRS V7 total column zone is also included in the comparison. The differences between the two algorithms on the ozone Degree of Freedom (DOF) reported in the L2 retrieval products and the channels used in the ozone retrieval are also reported.

5.2 Summary of Major Findings

Fig. 5.2.1 and Fig. 5.2.2 show the total ozone column as reported by CLIMCAPS-*Aqua* (left column), and compares the relative bias ([AIRS – OMI] / OMI) in total column for January 1, 2011 and the July 1, 2011, respectively. Biases of CLIMCAPS-*Aqua* (middle column) and AIRS V7 retrievals (right column) are shown for both the IR+MW (top) and IR-only (bottom) algorithms. All retrievals produce a similar spatial distribution of the total ozone with biases within ±10% from OMI. CLIMCAPS-*Aqua* shows a zonal, negative bias along the Southern Ocean for both IR+MW and IR-only in January, while AIRS V7 shows a positive bias in the northern high latitudes in July, especially over Siberia. These differences are also seen from the zonal mean of the biases shown in Figure 5.2.3 (top).

Note that the AIRS V7 ozone a priori come from an adjusted climatology (Yue and Lambrigtsen 2019), while the CLIMCAPS a priori are from MERRA2 (Smith and Barnet, 2020). In addition, the two retrieval systems calculate DOF differently as discussed in Section 9 of this report. The zonal averages of the DOF of signal and the number of AIRS-OMI matched observations used in calculating the averages are also shown in Fig. 5.2.3. Table 5.2.1 summarizes the global mean values for the two days of data. CLIMCAPS-*Aqua* and V7 relative biases against OMI are within a few percent of each other, and have similar yields. AIRS V7 reports significantly higher DOFs than CLIMCAPS-*Aqua*, particularly in the winter hemisphere. Based on the current analyses, this is potentially due to two differences in the retrieval algorithms: (a) AIRS V7 algorithm uses more basis functions than CLIMCAPS-*Aqua* for ozone (Fig. 5.2.4), and (b) AIRS V7 uses more AIRS spectral channels in the 10 µm ozone band (Fig. 5.2.5). The additional basis functions coupled with channels near the peak ozone absorption at 1040 cm^{-1} may provide additional resolution (and information) in the cold temperatures of the lower stratosphere near the tropopause.

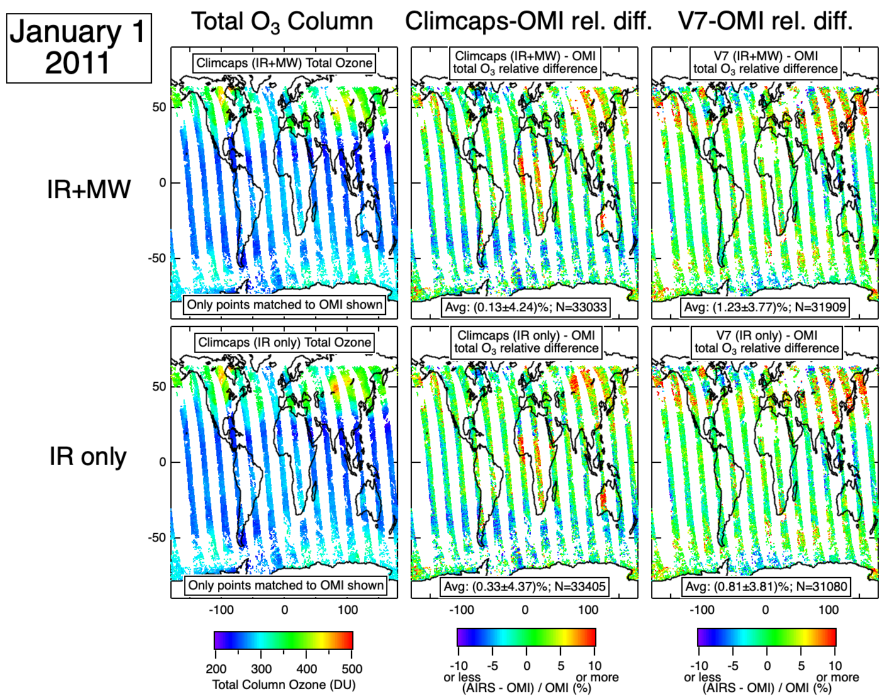


Figure 5.2.1. From left to right shows the CLIMCAPS-*Aqua* total column ozone, and relative bias for ClIMCAPS-*Aqua* and AIRS-V7 against OMI for January 1, 2011. Both IR+MW (upper) and IR-only (bottom) retrievals are shown. Relative bias is calculated as ([AIRS – OMI] / OMI) in percent. Data with QC=0 or 1 are used.

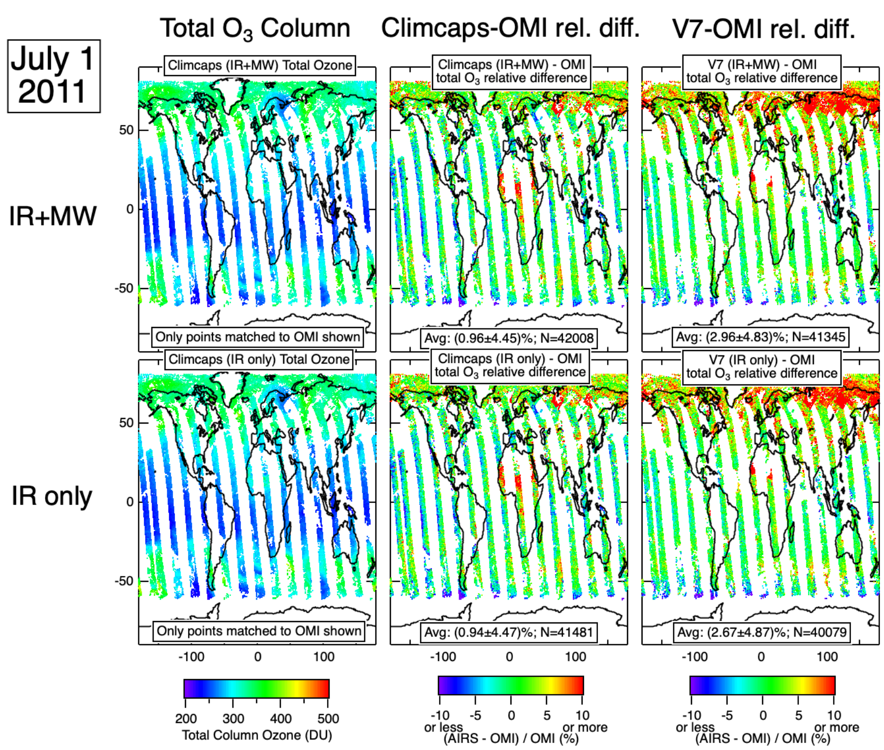


Figure 5.2.2. Similar to Fig. 5.2.1, but using data from July 1, 2011.

Table 5.2.1: Summary of average relative biases and ancillary results for CLIMCAPS-*Aqua* and AIRS V7 results against OMI total ozone columns

Date	Retrieval Version		Average (AIRS-OMI)/OMI bias (% ± 1σ std. dev)	No of matched observations	Average DOFS
Jan 1, 2011	CLIMCAPS-*Aqua*	IR+MW	0.13 ± 4.24	33033	1.93 ± 0.26
		IR-only	0.33 ± 4.37	33405	1.93 ± 0.26
	AIRS V7	IR+MW	1.23 ± 3.77	31909	2.59 ± 0.69
		IR-only	0.81 ± 3.81	31080	2.71 ± 0.70
July 1, 2011	CLIMCAPS-*Aqua*	IR+MW	0.96 ± 4.45	42008	2.03 ± 0.27
		IR-only	0.94 ± 4.47	41481	2.04 ± 0.27
	AIRS V7	IR+MW	2.96 ± 4.83	41345	2.70 ± 0.66
		IR-only	2.67 ± 4.87	40079	2.69 ± 0.67

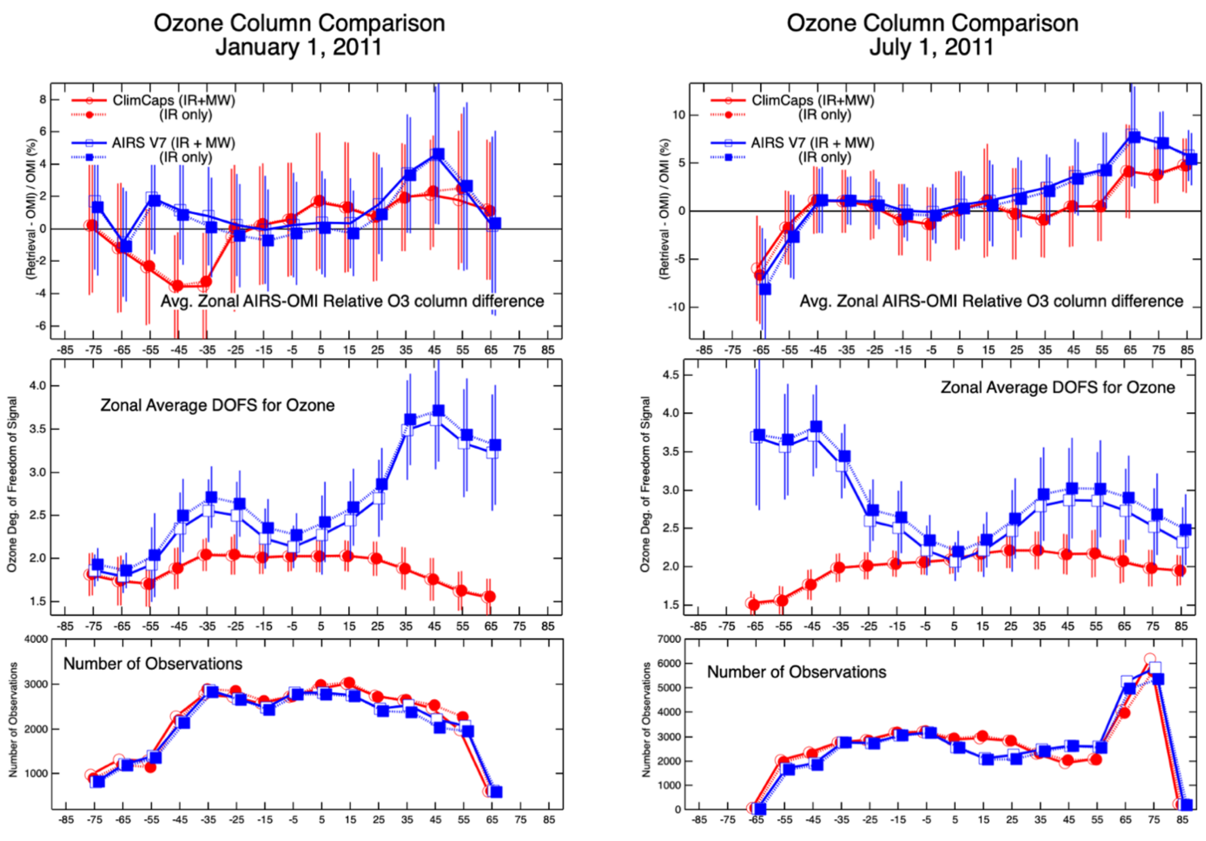


Figure 5.2.3. Zonally-averaged AIRS ozone bias relative to OMI (in 10° bins), degrees-of-freedom-of-signal, and number of observations matched to OMI. The left column is for Jan. 1, 2011, while the right column is for July 1, 2011. Note that the vertical axes can differ between the two dates.

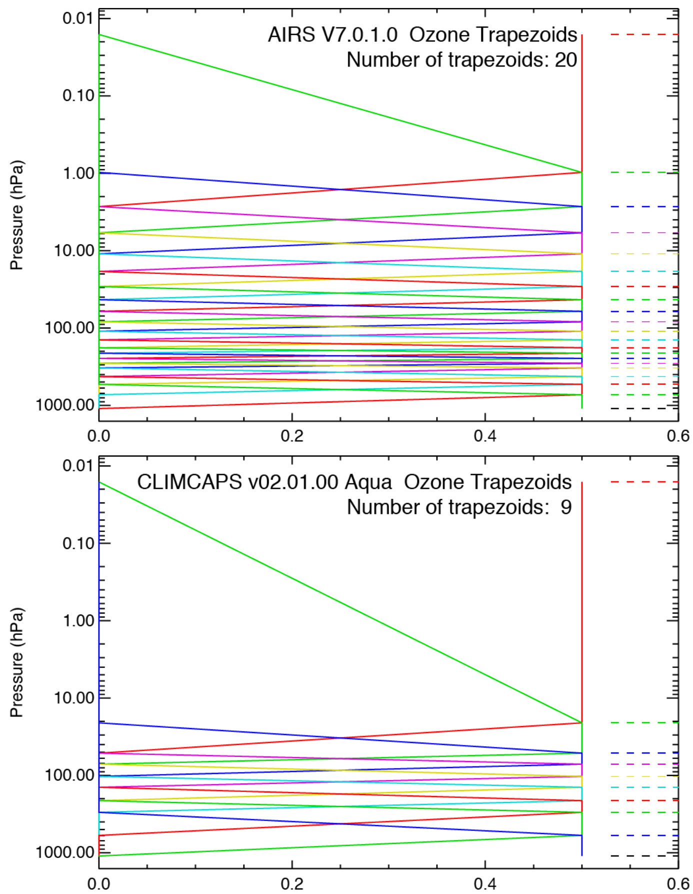


Figure 5.2.4. The 20 trapezoidal basis functions for ozone retrievals for AIRS V7 (upper panel) and the 9 functions used in CLIMCAPS-*Aqua* (lower panel).

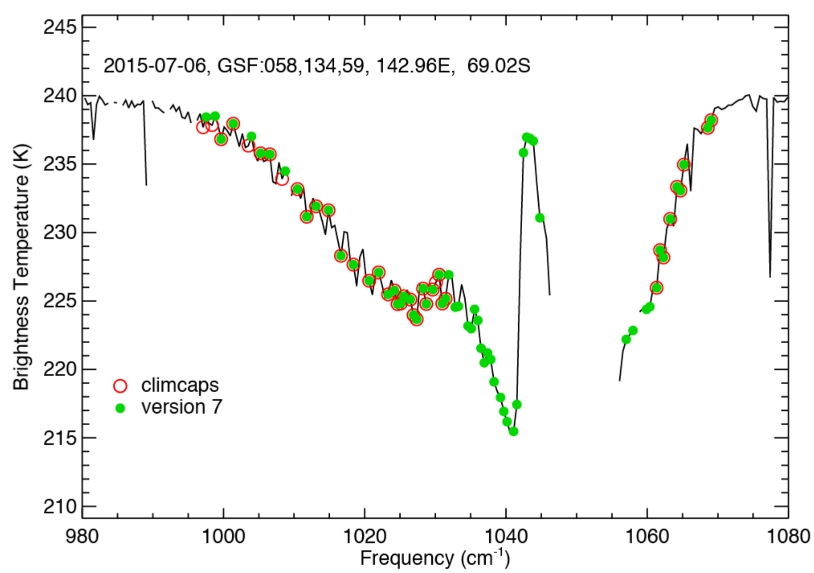


Figure 5.2.5. Spectral channels for AIRS ozone retrieval in V2 CLIMCAPS-*Aqua* (red open circles) and AIRS V7 (green closed circles) algorithms.

6. L2 Near Surface Air Temperature, Water Vapor, and VPD

6.1 Summary of Testing Tasks

L2 surf_air_temp (near surface air temperature, henceforth NSAT), L2 surf_spec_hum (specific humidity in units of g/kg, henceforth NSQ), surf_rel_hum (henceforth NSRH), and Vapor Pressure Deficit (VPD) from CLIMCAPS-*Aqua* IR-only and IR+MW retrievals are examined by comparing with the collocated in-situ measurements from MesoWest over land and International Comprehensive Ocean-Atmospheric Data Set (ICOADS) over ocean for July 2011.

The following collocation criteria are applied to match the satellite and in-situ observations: MesoWest observations obtained 15 min before or after each *Aqua* pass and within 13km from AIRS pixel; ICOADS observations obtained 1 hour before or after each *Aqua* pass and within 100 km from AIRS pixel. The spatially nearest data pair is selected if both constraints satisfied. To analyze the data-pairs, the quality control is performed using the quality flag from the IR+MW products. Any matchups with data quality marked as 2 are discarded.

6.2 Summary of Major Findings

6.2.1 Near Surface Air Temperature

Based on the results using July 2011 data, CLIMCAPS-*Aqua* retrievals agree very well with MesoWest measurements over land and with ICOADS measurements over ocean. Figure 6.2.1.1 left panels show that CLIMCAPS-*Aqua* NSAT tends to be colder than the MesoWest observations for high air temperature (above ~304 K). Over ocean, CLIMCAPS-*Aqua* NSAT is highly correlated with that from ICOADS (CC=0.92) with a consistent cold bias of -0.85 K. The RMSE of CLIMCAPS-*Aqua* NSAT is smaller over ocean than that over land. IR only and IR+MW comparison results are slightly different. But IR+MW has higher yields than IR only measurements (not shown).

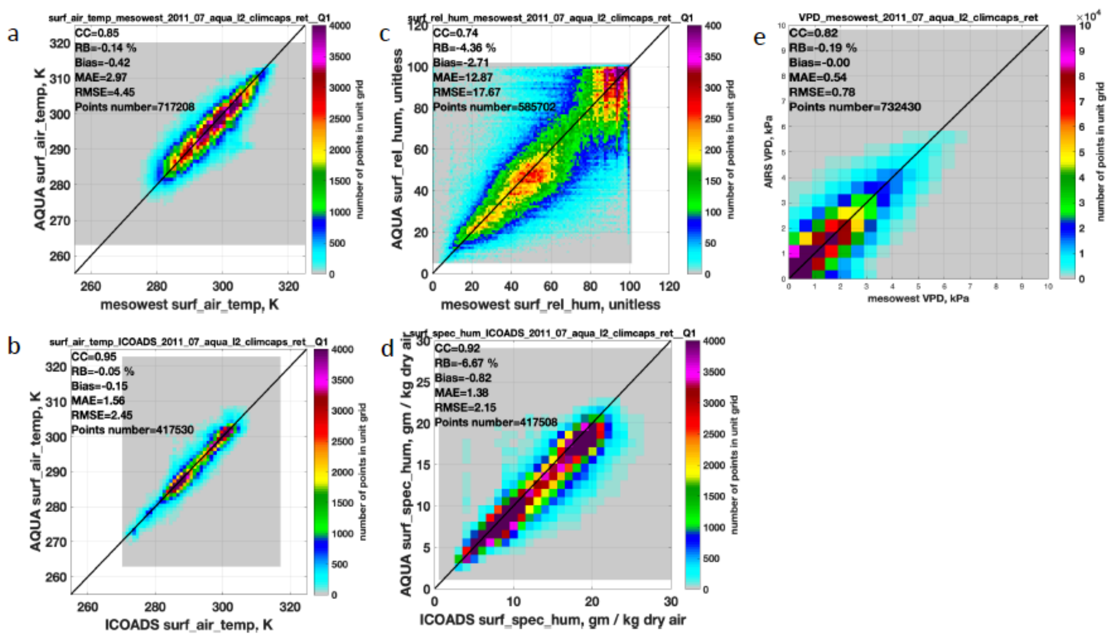


Figure 6.2.1.1 Scatterplots with colored data density of near surface air temperature (left, a and b), near surface moisture (middle, c and d), and vapor pressure deficit (VPD, right, e) retrieved by CLIMCAPS-*Aqua* IR+MW algorithm comparing observations from MesoWest (upper) and ICOADS (lower). Only MesoWest comparison is shown for VPD.

Fig. 6.2.1.2 shows the maps of bias and the RMSE between CLIMCAPS-*Aqua* NSAT evaluated using collocated MesoWest measurements and ICOADS data. Only the IR+MW retrieval is shown since very small differences are seen between IR+MW and IR-only CLIMCAPS-*Aqua* near surface retrievals. The matched points have been put into 2°x2° grid. The bias of NSAT from CLIMCAPS-*Aqua* is generally within 2 K compared to in-situ observations. The Great Plain region (e.g. Oklahoma, Texas, Kansas) compares the best with the bias less than 0.1 K. For areas in high altitude, the retrieval shows a larger bias. RMSE also shows similar pattern. The Great Plain region has the lowest RMSE (less than 2 K) while the Northwest region has a higher RMSE of ~5 K. Over ocean, the bias is smallest in the Atlantic Ocean region, ~-0.5 K. The RMSE is also the lowest in this region.

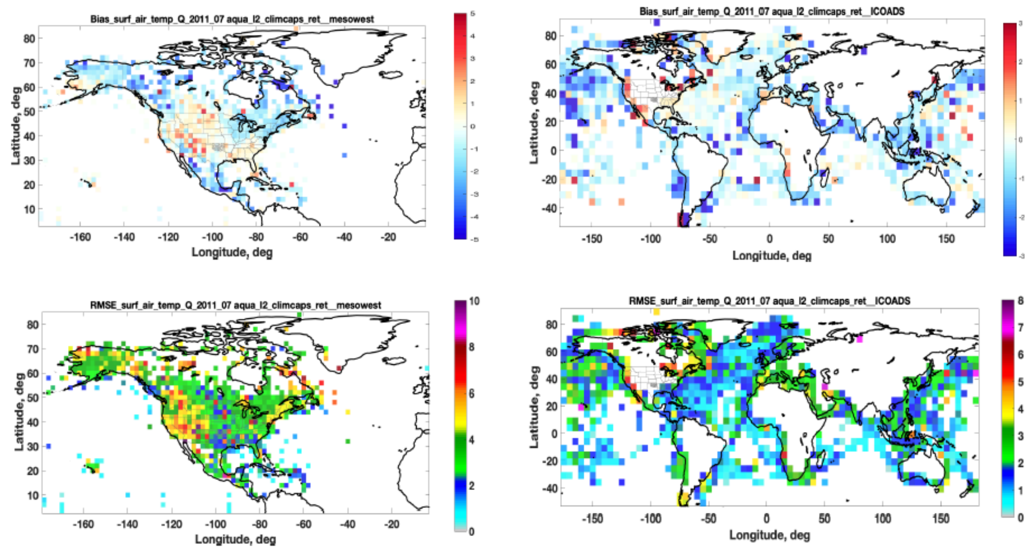


Figure 6.2.1.2 The mean bias (upper) and RMSE (lower) for CLIMCAPS-*Aqua* IR+MW near surface air temperature retrievals calculated against collocated MesoWest (left) and ICOADS (right). Unit is in K for both bias and RMSE.

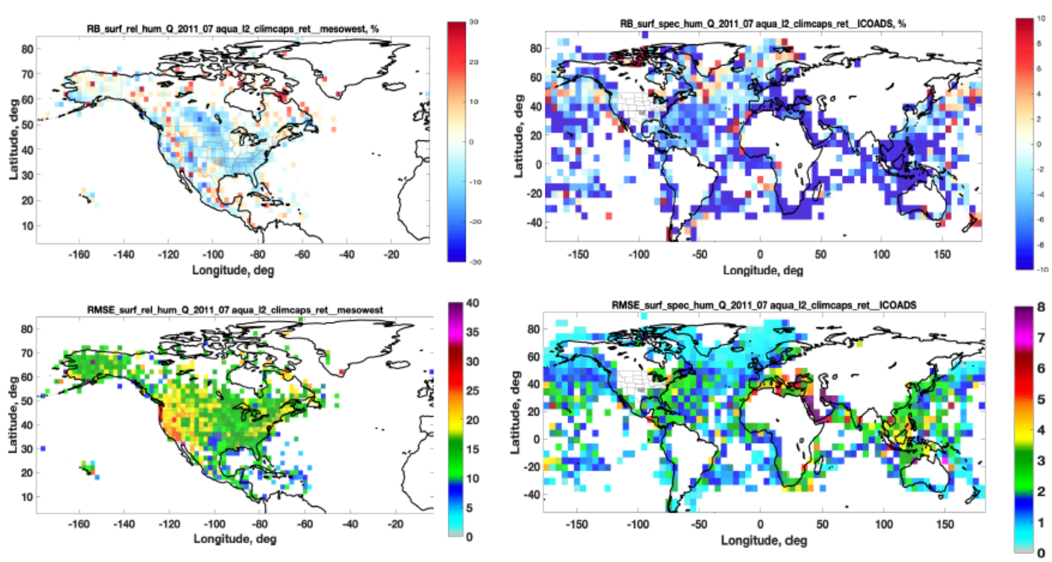


Figure 6.2.2.1 The mean relative bias (upper) and RMSE (lower) for CLIMCAPS-*Aqua* IR+MW near surface humidity retrievals calculated against collocated MesoWest (left) and ICOADS (right). Near surface relative

humidity (%) is evaluated using the MesoWest data, and the near surface specific humidity (g/kg) is evaluated using the ICOADS data. Note that humidity RMSE is not normalized by mean climatology.

6.2.2 Near Surface Humidity and Vapor Pressure Deficit

CLIMCAPS-*Aqua* NSQ and NSRH agree well with in-situ measurement. Although the relative mean bias is small (-4%) over land, Fig. 6.2.1.1c and d show a wide range of scatter, especially at RH >85%. CLIMCAPS-*Aqua* NSRH is highly correlated with ICOADS (CC=0.92) and the mean relative bias is ~-7%. Similar to NSAT, the results of humidity from IR-only and IR+MW differ slightly with a higher yield from IR+MW (not shown).

Fig. 6.2.2.1 show the maps of relative bias (%) and the RMSE (%) between CLIMCAPS-*Aqua* NSRH calculated from MesoWest and for NSRH relative bias (%) and RMSE (g/kg) of NSQ compared with ICOADS data. The relative bias of NSQ is within 20%. The East part of the United States has a consistent dry bias ~10% and a RMSE < 20%, while the Western mountainous United States has a larger magnitude of bias and RMSE reaching 30%. Over Ocean, a dry bias is generally observed.

Vapor Pressure Deficit (VPD) is an important application of CLIMCAPS-*Aqua* near surface air temperature and water vapor on drought early warning. The CLIMCAPS-*Aqua* Level 2 VPD product compares very well with MesoWest observations as shown in Fig. 6.2.1.1e.

7. Total Column Water Vapor (TCWV)

7.1 Summary of Testing Tasks

The V2 CLIMCAPS-*Aqua* total column water vapor (TCWV) retrievals are evaluated by comparisons with pixel-scale collocated measurements from the North America SuomiNet ground-based GPS stations over land and the monthly observations over global ocean by the Advanced Microwave Scanning Radiometer - Earth Observing System (AMSR-E) on board of *Aqua* obtained from Remote Sensing System (Fig. 7.1.1).

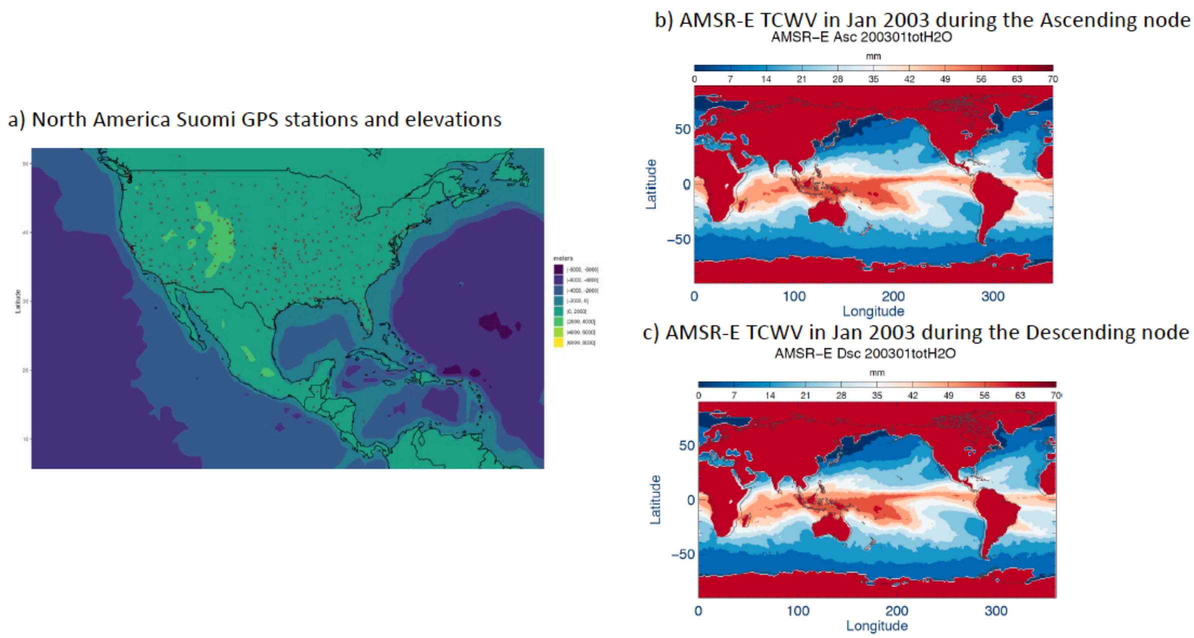


Figure 7.1.1 a) Elevation map with location of SuomiNet GPS stations (marked with an X). AMSR-E monthly mean total column water vapor (mm) in Jan 2003 for ascending (b) and descending (c) nodes separately.

Matchups between AIRS data and the SuomiNet GPS stations were made following a similar analysis to that of Roman et al. (2016). Matchups had to be within 1 hour and 1° radius of one another. Since TCWV depends on the vertical dimension of water vapor, matchups needed to be within 100 meters of one another to ensure comparable comparisons. The 10 closest AIRS FOVs to the station that meet these temporal, spatial and vertical requirements are averaged together to create the matched TCWV value. SuomiNet TCWV values in which the error (pwv_err from the data) was greater than 5% were discarded. Results are presented using relative bias (%) with a 95% uncertainty estimate of the mean included as an error bar. AIRS retrieval QC for each individual retrieval systems is applied and retrievals with QC=0 or 1 are included in the comparison.

For AIRS and AMSR-E comparisons, daily gridded TCWV in the Ascending (Asc) and Descending (Dsc) passes from the two instruments are used to examine the difference between the day- and night-time AIRS retrievals over ocean. Comparisons with other microwave-based TCWV products have been conducted, such as SSMI, GMI, and TMI. Similar results are obtained with the AMSR-E comparisons reported here. Both AIRS and AMSR-E are on board NASA's *Aqua* satellites, which ensures the collocation of the measurements. Monthly mean differences between the AMSR-E and the AIRS TCWV data are calculated for Asc and Dsc passes separately. Results are presented on maps of 1×1° longitude by latitude grid boxes.

7.2 Summary of Major Findings

7.2.1 Yield of TCWV Retrieval

Figure 7.2.1 shows the yield of the total column water vapor from CLIMCAPS-*Aqua* IR+MW retrievals for ascending and descending nodes separately using one month of data (January 2003). Comparing to AIRS V7, CLIMCAPS-*Aqua* has larger yields over high latitude region and subtropical low cloud covered regions. Both CLIMCAPS-*Aqua* and AIRS V7 retrievals have larger daytime yield over subtropical low cloud covered regions with CLIMCAPS-*Aqua* producing smaller day-night differences. However, additional day-night sampling differences are seen in CLIMCAP-*Aqua*, which has a smaller daytime yield over certain land regions such as Australia and a larger nighttime yield over Europe and central Asia, but CLIMCAPS-*Aqua* does not have the sharp day-night difference along 60ºN latitude line.

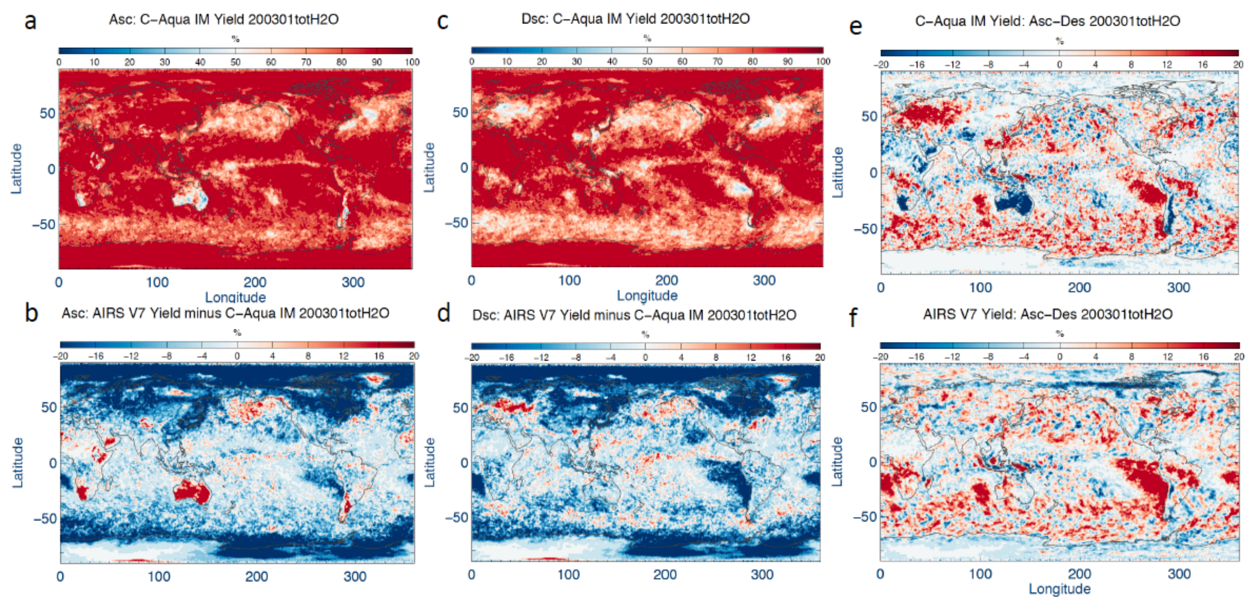


Figure 7.2.1 CLIMCAP-*Aqua* IR+MW total column water vapor retrieval yield in January 2003 for the ascending (a) and descending (c) nodes and their differences with AIRS V7 IR+MW (b for ascending and d for descending). The yield differences between the two nodes (ascending-descending) for CLIMCAPS-*Aqua* and AIRS V7 IR+MW retrievals are shown in e) and f) respectively.

7.2.2 L2 TCWV over Land

As shown in Fig. 7.2.2.1, negligible differences on TCWV retrievals between CLIMCAPS-*Aqua* IR-Only and IR+MW retrieval systems are seen. All different AIRS retrievals show similar magnitude of biases, which in general is a wet bias in dry conditions (TCWV< 20 mm) and a dry bias in wet conditions (TCWV>60mm). All retrieval systems have a ±5% bias for intermediate TCWV values. Note than for comparison purpose, QC=0 and QC=2 results are also shown although there is sampling difference for QC=0 for AIRS V7 and CLIMCAPS-*Aqua* systems (see section 4.2.1) and QC=2 cases are recommended as "not to use" retrievals. Fig. 7.2.2.2 shows the histogram of the bias. CLIMCAPS-*Aqua* has a much narrower error histogram that is centered around 0 in January than in July suggesting a better agreement between CLIMCAPS-*Aqua* in January. Moreover, the January density error histogram from CLIMCAPS-*Aqua* is narrower than AIRS V7 while similar distributions are seen in July between the two. The relationship between the error distribution and the background condition is also investigated by the error histogram in different surface relative humidity bins as measured by the GPS stations (Fig. 7.2.2.3). Both CLIMCAPS-*Aqua* and AIRS V7 retrievals have a widening of the error histogram with increasing surface relative humidity. This suggests the range of error increases with increasing surface relative humidity. In addition, the peak shifts from the right to the center (decreasing magnitude of wet bias) as surface relative humidity increases.

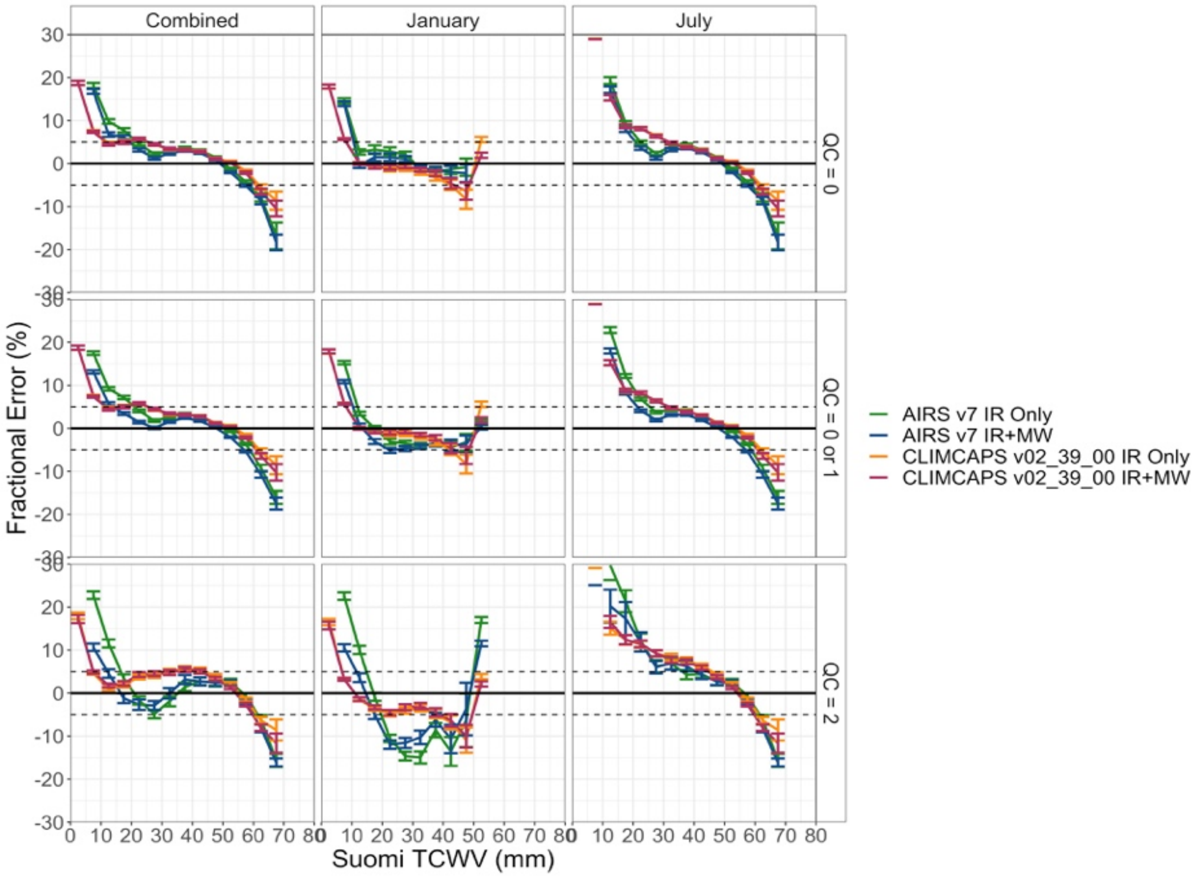


Figure 7.2.2.1 Mean relative bias in % by Suomi TCWV by retrieval systems (color), months (columns) and QC (rows). Error bars represent the 95% uncertainty estimate of the mean. Data in January and July of 2011 are shown in the figure.

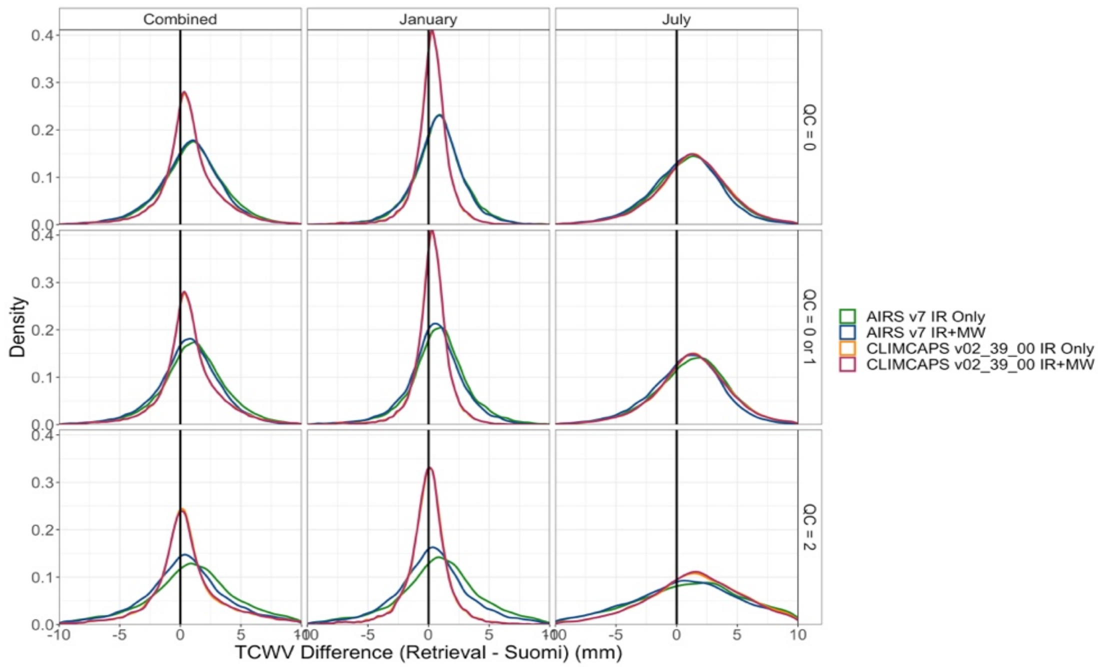


Figure 7.2.2.2 Histogram of mean bias magnitude (mm, AIRS-GPS) for AIRS TCWV by retrieval systems (color), months (columns) and QC (rows). Data in January and July of 2011 are shown in the figure.

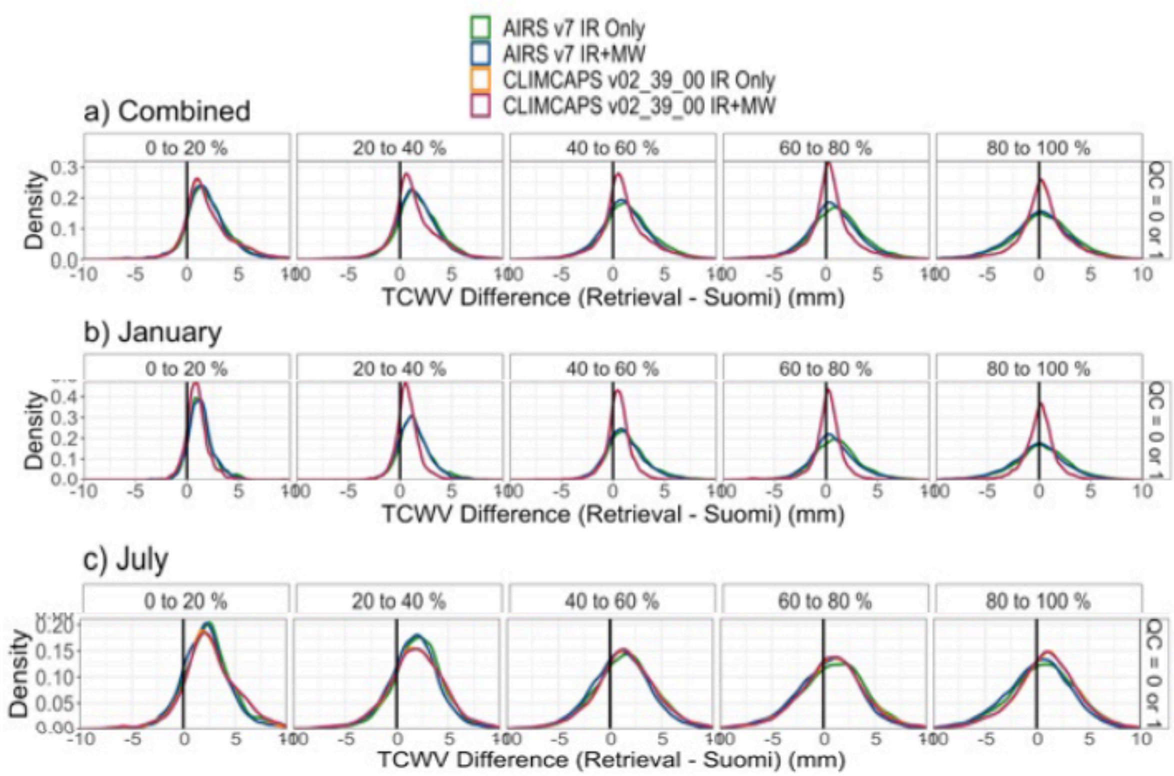


Figure 7.2.2.3 Error density histograms (Retrieval - Suomi) by Surface Relative Humidity (columns), and retrieval system (color).

7.2.3 Monthly Mean TCWV over Ocean

Figure 7.2.3.1 summarizes the bias of various AIRS TCWV retrievals evaluated using the AMSR-E measurements. Differences between the CLIMCAP-*Aqua* IR+MW and IR-only retrievals are neglectable, therefore, only its IR+MW results are shown. In general, CLIMCAPS-*Aqua* has a dry bias over ocean which peaks at the deep convective zones such as ITCZ and midlatitude storm track regions. AIRS V7 also shows a dry bias associated with deep convection, but a wet bias is seen over the regions with shallow convection and low clouds. Therefore, AIRS V7 TCWV bias shows a stronger cloud type dependent and regional variations than CLIMCAPS-*Aqua*. Comparing the monthly mean TCWV from different retrieval systems, CLIMCAPS-*Aqua* has a smaller dry bias than AIRS V7 over deep convective regions over ocean. CLIMCAPS-*Aqua* is wetter than AIRS V7 over the Amazon and Australia, which requires further validation over land with other reference datasets. CLIMCAPS-*Aqua* is wetter than AIRS V7 over shallow cloud covered oceanic regions and certain land regions including S.E Asia, Sahal, E. China and E. United States. The differences between AIRS V7 IR-only and CLIMCAPS-*Aqua* have similar spatial patterns to those between AIRS V7 IR+MW and CLIMCAPS-*Aqua* but with larger magnitudes.

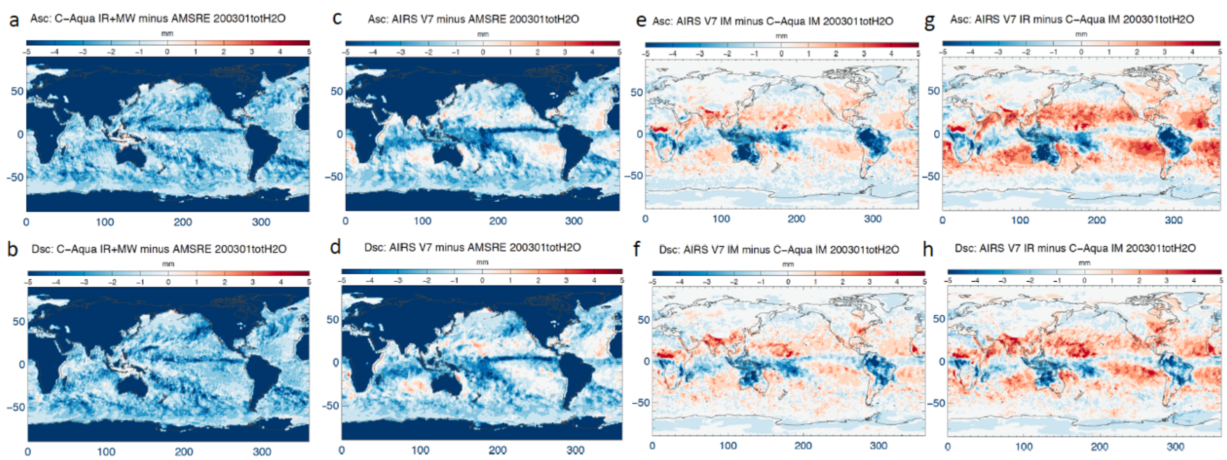


Figure 7.2.3.1 The monthly mean biases of TCWV in CLIMCAP-*Aqua* IR+MW (a and b) and AIRS V7 IR+MW (c and d) by comparing with AMSR-E (see Fig. 7.1.1) for ascending (a and c) and descending (b and d) passes. The TCWV differences between various AIRS retrievals are shown: differences between AIRS V7 IR+MW and CLIMCAPS-*Aqua* for ascending (e) and descending (f) passes; differences between AIRS V7 IR-only and CLIMCAPS-*Aqua* for ascending (g) and descending (h) passes. Data from January 2003 is used in the calculation.

8. Cloud Parameters

The CLIMCAPS-*Aqua* L2 retrievals of two-layer effective cloud fraction (ECF) and cloud top temperature (TCldTop) are compared with the AIRS V6 cloud retrievals, which have been extensively evaluated against MODIS, CloudSat, and CALIPSO measurements (e.g. Kahn et al. 2014). Comparisons are conducted at both the L2 pixel scale and the gridded average (L3). No quality control was applied as this report aims to understand all of the potential artifacts in the cloud products within selected AIRS granules. These cloud products are used in a wide variety of geophysical states, including temperature and specific humidity soundings, and the quality control of the retrievals.

8.1 L2 Cloud Retrievals

Three granules are chosen to demonstrate the differences between CLIMCAPS-*Aqua* and AIRS V6 cloud retrievals: granules 010, 055, and 121 on January 1, 2016. Granule 010 is situated along the subtropical west African coast, granule 055 is centered along the eastern side of Greenland, and granule 121 extends from northern Alaska into the Canadian Arctic archipelago. These three granules extend across a series of cloud regimes in the low and high latitudes that span most surface classes and cloud types. Granule 010 captures subtropical oceanic stratocumulus and shallow cumulus to the south and west, and tropical land-based deep convection to the north and east with some scattered thin cirrus. Granule 055 captures two storm systems with a relatively clear area in between over sea ice adjacent to Greenland. Granule 121 captures complex clouds over a wide variety of ice surfaces. Their brightness temperature at 1231cm^{-1} and surface classification are shown in Fig. 8.1.1. There are some small changes in the MW surface classification between the two algorithms, which tend to occur for very few FOVs at the edge of regions with different surface classes. As a result, we can have confidence that discrepancies in surface classification appears to play little role in between v6 and CLIMCAPS-*Aqua* cloud products.

The ECF and TCldTop for the three granules are presented in Figs. 8.1.2-4, which shows the following differences between the CLIMCAPS-*Aqua* and AIRS V6:

1. The total ECFs from the two retrievals largely agree with each other on pixel level.
2. Larger differences are seen over the subtropical low cloud covered oceanic granule. CLIMCAPS-*Aqua* appears to have more difficulty in retrieving physically reasonable values of ECF and TCldTop for certain cloud types compared to v6, in particular shallow cumulus and stratocumulus. The ECF in V6 is near 1.0 most everywhere over the stratocumulus cloud deck with a few occasional cloud retrieval failures seen as empty squares as in Granule 010. This behavior is much more prevalent in CLIMCAPS-*Aqua*, which produces many more pixels with ECF significantly less than 1.0. This is consistent with the behavior seen in the cloud top temperature, where lower values of ECF coincide with lower values of cloud top temperature (higher cloud top). This suggests that both V6 and CLIMCAPS-*Aqua* are fitting the observed radiances well, but CLIMCAPS-*Aqua* tends toward lower values of ECF and cloud top temperature than is thought realistic for subtropical stratocumulus. This is reminiscent of problems that existed in V5 and earlier versions of the AIRS Team algorithm. Comparisons betweenV5 and V6 are shown in Kahn et al. (2014), and V5 tends to place clouds too high with a reduced ECF.
3. There are differences in the clouds over some ice surfaces especially in scenes thought to have little or no cloud cover. CLIMCAPS-*Aqua* contains spotty cloudiness over some of these ice surfaces while V6 does not. Given these differences, more investigations are necessary to

determine which algorithm produces better cloud retrievals over snow and ice covered surfaces. The total ECF fields within the two Arctic granules 055 and 121 are complex and varied. The horizontally extensive and thick clouds associated with baroclinic storm systems generally appear more similar between the two retrievals. However, in the thinner, shallow, warmer, and broken clouds, there are significant differences at the pixel-scale as well as at sub-synoptic scales. In granule 055, note the strong differences in ECF over the sea ice adjacent to the eastern edge of Greenland. The sea ice patterns are nearly identical between v6 and CLIMCAPS-*Aqua* (Figure 8.1.1). Therefore, the differences do not arise from the very minor changes in surface classification.

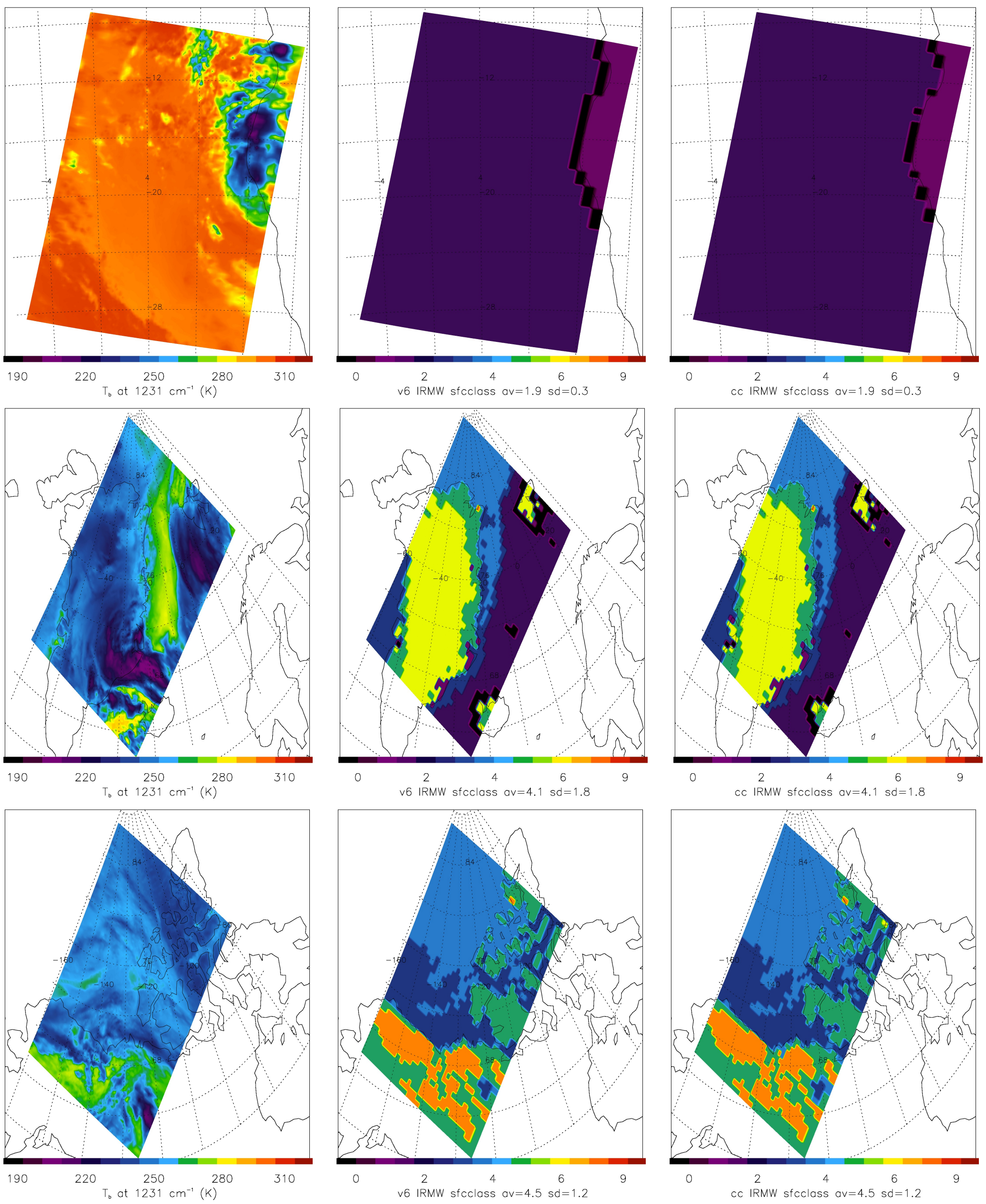


Figure 8.1.1. Upper row: granule 010 1231 Tb (k) (left), v6 MW surface classification (middle), and CLIMCAPS-*Aqua* surface classification (right). Middle row: same as above except for granule 055. Lower row: same as above except for granule 121. Coastline=black. Glacier ice=yellow. Low emissivity

snow=orange. High emissivity snow=green. Low emissivity ice=light blue. High emissivity sea ice=dark blue. Open ocean=dark purple. Land=light purple. All three granules observed on January 1, 2016.

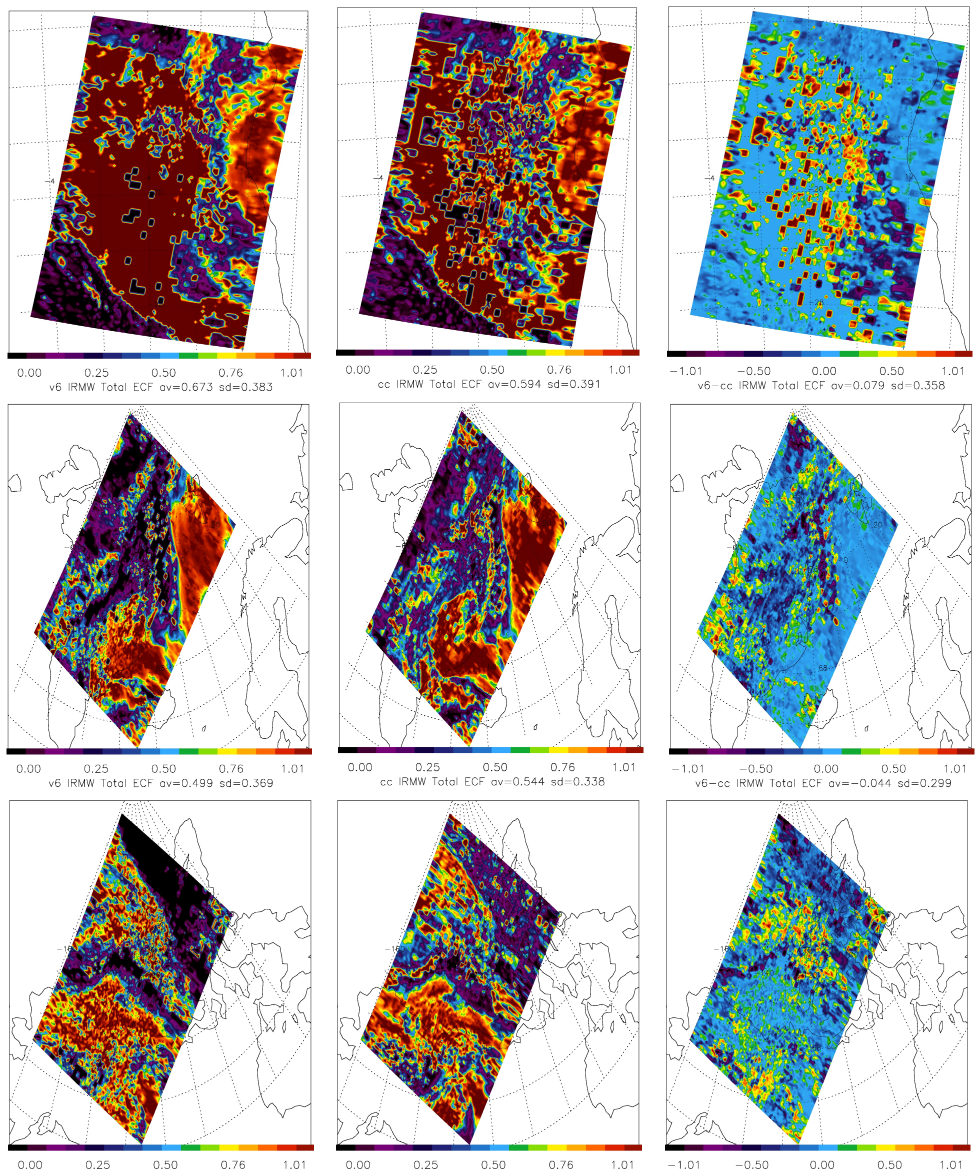


Figure 8.1.2. Total two-layer effective cloud fraction (ECF) for granules 010 (upper row), 055 (middle row), and 121 (lower row) for v6 (left column), CLIMCAPS-*Aqua* (middle column), and the difference between v6-CC (right column). The average ECF and its standard deviation are reported over the full granules in the left and middle columns. The average difference and its standard deviation are reported in the right column.

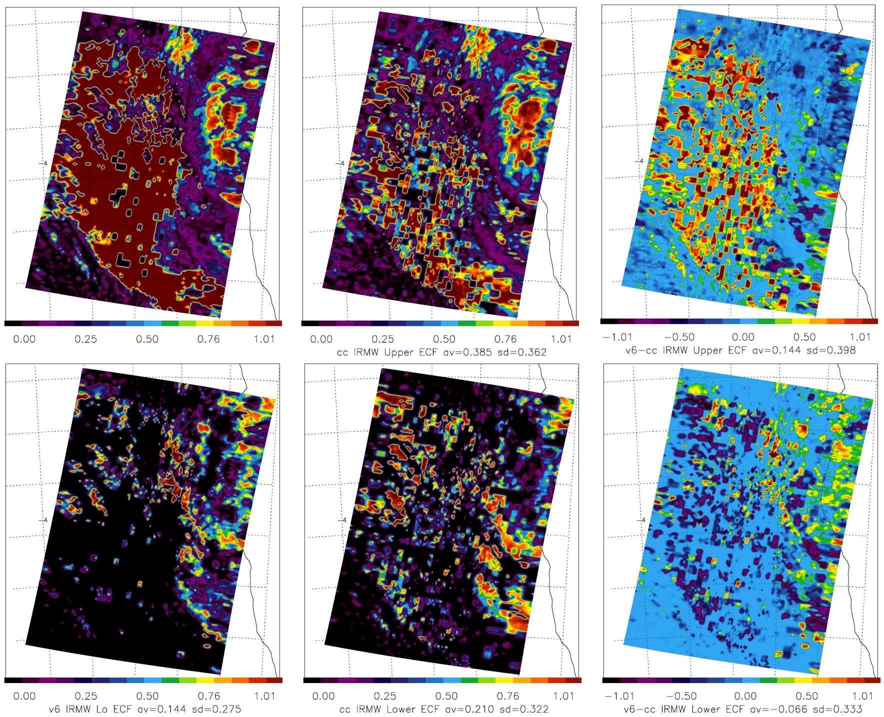


Figure 8.1.3. Upper row: upper layer ECF and lower row: lower layer ECF for granule 010 for v6 (left column), CC (middle column), and the difference of v6 - CLIMCAPS-*Aqua* (right column).

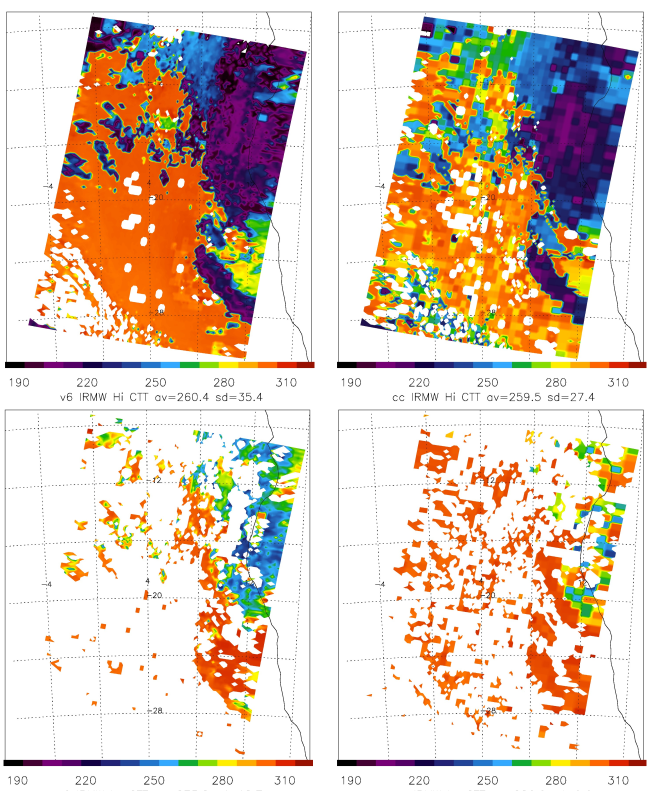


Figure 8.1.4. Upper row: upper layer TCldTop and lower row: lower layer TCldTop for granule 010 for v6 (left column) and CLIMCAPS-*Aqua* (right column).

8.2 L3 Cloud Products

The differences on L3 cloud products are consistent with that of L2 findings. All data shown are obtained from one month of L2 granules using the IR-only retrievals. The granules are then gridded to 1x1 degree resolution over the entire globe. While only one month is shown in the report, a series of spot checks were done over a few other months and the results obtained for other time periods are generally similar in behavior with expected variability over the seasonal cycle. The variables evaluated are the following: the daily (ascending plus descending orbits) total and layer ECF and cloud top temperature (Tcld) in Fig. 8.2.1 and 8.2.2, respectively.

1. The global averaged differences in total ECF between V6 and CLIMCAPS-*Aqua* are less than -0.02, with 0.424 global mean reported by V6 and 0.438 reported by CLIMCAPS-*Aqua*. The regional differences, however, are larger and frequently change sign depending on latitude, the presence of land, ocean, sea or land ice, and cloud regimes such as tropical deep convection/cirrus and subtropical stratocumulus/shallow cumulus. For instance, the differences are lowest over the tropical oceans while they are much larger over land. The total ECF is generally larger for V6 in the tropical land areas while CLIMCAPS-*Aqua* is larger over some higher latitude land masses and especially over the Arctic region. The stronger contrasts in ECF along the coasts in the tropics in V6 suggest some dependence of ECF on surface type. Further investigation is warranted. There is a hint of some striping in the difference plot (upper right of Fig. 8.2.1) which is attributed to the V6 total ECF. There is some scan angle dependence of total ECF in V6, while this appears to be less the case in CLIMCAPS-*Aqua*. While subtle, this behavior is also seen in the upper Tcld (Fig. 8.2.2).
2. The global mean value of upper level ECF (middle row of Fig. 8.2.1) is 0.265 for V6 and 0.312 for CLIMCAPS-*Aqua*, for a difference of approximately -0.05. The global mean value of lower level ECF (lower row of Fig. 8.2.1) is 0.159 for V6 and 0.126 for CLIMCAPS-*Aqua*, a difference of -0.033. The two retrievals behave much differently with regard to whether cloud is placed in the upper or lower layers with regional differences. There is more cloudiness placed into the upper layer in CLIMCAPS-*Aqua* within the high latitude polar regions and additionally within the deep tropics. However, there is more cloudiness placed into the upper layer of V6 within the subtropics, especially in the stratocumulus cloud regimes near Peru and Namibia.
3. The global averaged differences between V6 and CLIMCAPS-*Aqua* for the upper and lower layers of Tcld are -1.1 and -6.9 K for the upper and lower layers, respectively. While the spatial patterns show some similarity, one can see that the Tcld is much colder over the Peruvian and Namibian stratocumulus regions in CLIMCAPS-*Aqua*. This behavior also appears in shallow subtropical cumulus throughout the low latitudes surrounding areas of deeper tropical convection. In the heart of the deep convection in the central Pacific, V6 has upper level Tcld 5-10 K warmer than CLIMCAPS-*Aqua*. Many of these clouds are thin cirrus and it could be possible that CLIMCAPS-*Aqua* is better able to assign a correct altitude to these clouds. The global mean value of lower level Tcld (lower row of Fig. 8.2.2) is 262.4 K for V6 and 268.2 K for CLIMCAPS-*Aqua*, a difference of nearly -5.8 K. One can see that the lower layer cloud in deep convection is placed at a higher altitude (colder Tcld) in V6. In the subsidence regions, the Tcld is fairly uniform and does not indicate much structure related to marine boundary layer depth or SSTs.
4. As Kahn et al., (2008) showed, within convective systems, the V6 algorithm will frequently assign higher values of ECF to a lower cloud layer if (i) the convective cloud top is diffuse, (ii) the anvil cloud is thin with discernible lower-layer clouds, and (iii) thin cirrus is occurring in

proximity to convection with discernible lower-layer clouds. Similar issues are at play in the NH and SH storm tracks where nimbostratus clouds are common (this cloud type has very diffuse tops upon examination of CloudSat reflectivity data). In the storm track regimes, it appears CLIMCAPS-*Aqua* assigns more ECF to the lower cloud layer than V6. Further investigation into the causes of when more or less ECF is apportioned to one layer or another is warranted.

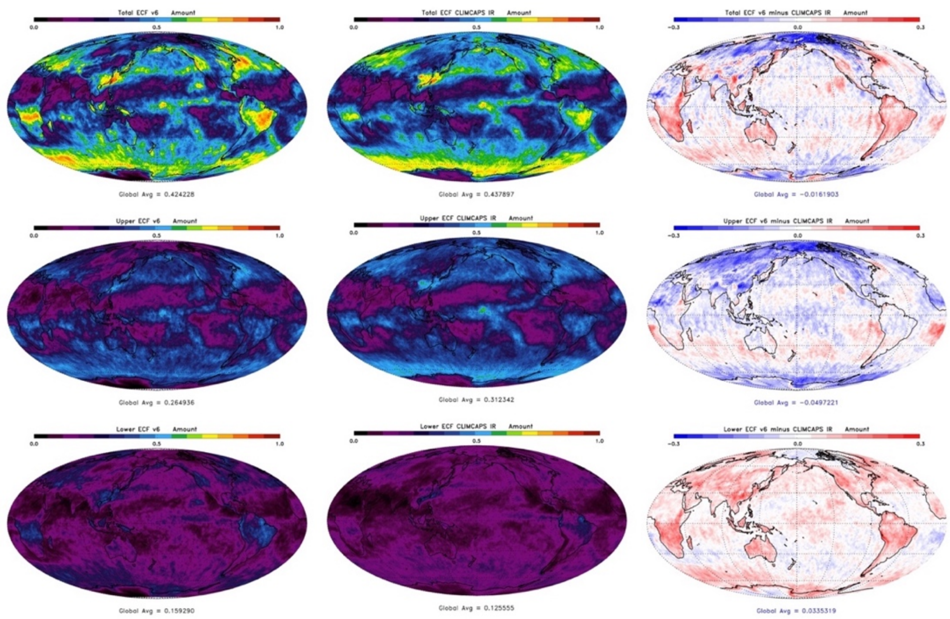


Figure 8.2.1. The effective cloud fraction (ECF) for v6 and CLIMCAPS-*Aqua* for January 2016 (in amount from 0.0 to 1.0). Upper row: total ECF (sum of upper and lower layers). Middle row: upper level ECF. Lower row: lower level ECF. Left column: v6. Middle column: CLIMCAPS-*Aqua*. Right column: difference of v6 minus CLIMCAPS-*Aqua*. Global average values are reported at the bottom of each panel.

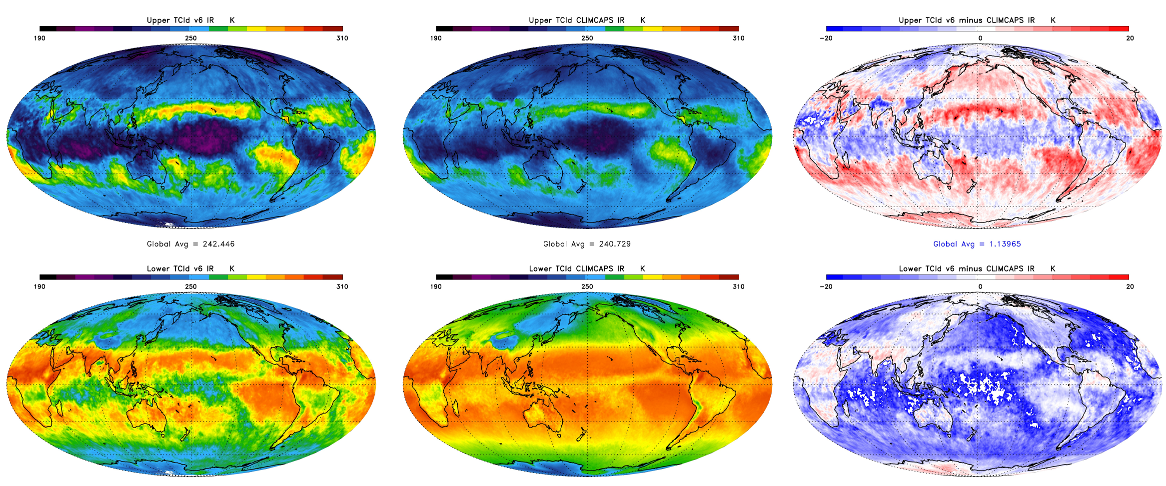


Figure 8.2.2. The cloud top temperature (Tcld) for v6 and CLIMCAPS-*Aqua* for January 2016 (in deg K). Upper row: upper level Tcld. Lower row: lower level Tcld. Left column: v6. Middle column: CLIMCAPS-*Aqua*. Right column: difference of v6 minus CLIMCAPS-*Aqua*. Global average values are reported at the bottom of each panel.

9. Degree of Freedom for Temperature and Water Vapor Profiles

9.1 Summary of Testing Tasks

The degrees-of-freedom (DOF) is the trace of the diagonal of the averaging kernel ($\boldsymbol{A}$, a matrix), and is reported in the L2 products of the AIRS V7 and CLIMCAPS-*Aqua* retrievals. The CLIMCAPS-*Aqua* and V7 DOF results for temperature and water vapor are compared using data from January 1, 2011, and July 1, 2011.

DOF is calculated as the sensitivity of changes in the retrieved state to changes in the "true" state (e.g., Rodgers, 2000). That is,

$$\boldsymbol{A} = \frac{\partial \hat{x}}{\partial \boldsymbol{x}} = \frac{\partial \hat{x}}{\partial \boldsymbol{F}} \frac{\partial \boldsymbol{F}}{\partial \boldsymbol{x}} = \boldsymbol{GK}, \tag{9.1}$$

where $\boldsymbol{F}$ is computed radiance for the retrieval state vectors using the forward model, the gain matrix $\boldsymbol{G} = \partial \hat{\boldsymbol{x}} / \partial \boldsymbol{F}$ is a measure of the sensitivity of the retrieval, $\hat{\boldsymbol{x}}$, to changes in the radiance, and the Jacobian $\boldsymbol{K} = \partial \boldsymbol{F} / \partial \boldsymbol{x}$ is the matrix of derivatives of the radiance to changes in each element of the state vector. The DOF is a useful metric of how much independent information is gotten from the retrieval.

It is worth noting that averaging kernels are calculated differently for AIRS V7 and CLIMCAPS-*Aqua*. The V7 retrieval algorithm does not use an a priori covariance as a constraint. Rather, a singular-value-decomposition (SVD) technique is used, with low-eigenvalued modes explicitly damped out (see Section III of Susskind et al., 2003); the DOF is the equivalent of the fractional number of significant eigenfunctions used (see Maddy and Barnet, 2008). CLIMCAPS-*Aqua* uses an optimal estimation technique in which the gain, $\boldsymbol{G}$, uses an a priori covariance (Smith and Barnet 2019, 2020). As discussed in Sec. 2.4.2 of Rodgers (2000), it can be shown that the DOF (when using an a priori covariance) is equivalent to the number of "effective" eigenfunctions in an SVD-transformed averaging kernel. Additionally, the a priori for CLIMCAPS-*Aqua* is MERRA2 reanalysis, while the a priori for V7 are from neural net retrievals (Milstein and Blackwell, 2015). For both CLIMCAPS-*Aqua* and AIRS V7, calculation of the DOF is from the "physical" retrieval, but V7 also gets *unquantified* information from the neural net which uses the AIRS radiance spectra. As noted in Sec 6.1.3 of Yue and Lambrigtsen (2019), the DOFs for V7 temperature and water vapor reflect the weighting between the neural-net and physical retrievals, thus the DOFs reported in the V7 L2 product may not represent the information content of the *entire* retrieval from the AIRS spectra. Therefore, the discussion below only intends to show the quantitative differences on the DOF for temperature and water vapor profiles reported in the L2 retrieval products. These differences are partially caused by the different methods to calculate averaging kernels and DOFs used in the AIRS V7 and CLIMCAPS-*Aqua* algorithms. In this study, the DOF data were included if at least part of a retrieved temperature or water vapor profile passed quality control. Different filtering methods have been tested and it is found that the differences between the DOFs from different algorithms are not caused by sampling.

9.2 Summary of Major Findings

Figure 9.2.1 compares daytime global maps of retrieval DOF for CLIMCAPS-*Aqua* and AIRS V7 IR+MW products for January 1, 2011. The DOFs for temperature (top row) have similar spatial patterns between CLIMCAPS-*Aqua* and V7, with V7 being slightly less than CLIMCAPS-*Aqua* on average. For water vapor (bottom row), V7 shows significantly lower DOFs, which reflects changes in the V7 algorithm that increased damping and removed shortwave channels used in the water vapor retrieval. These differences are highlighted in Figure 9.2.2 by comparing the zonal averages for January 1, 2011 and July 1, 2011. Both the IR+MW and IR-only retrievals of CLIMCAPS-*Aqua* and V7 are shown. The zonal averages in January and July both show about similar shape with offsets reflecting differences in the weighting given the cloud-cleared spectra in the physical retrieval, with a pronounced difference for water vapor between CLIMCAPS-*Aqua* and AIRS V7. The smaller water vapor DOF of AIRS V7 is the result of lower weighting given the spectra in the V7 physical retrieval as discussed in detail by Yue and Lambrigtsen (2019). The dips in temperature DOF (and water vapor DOF to a lesser extent) near the equators are likely from higher spectral noise because of clouds in the Intertropical Convergence Zone.

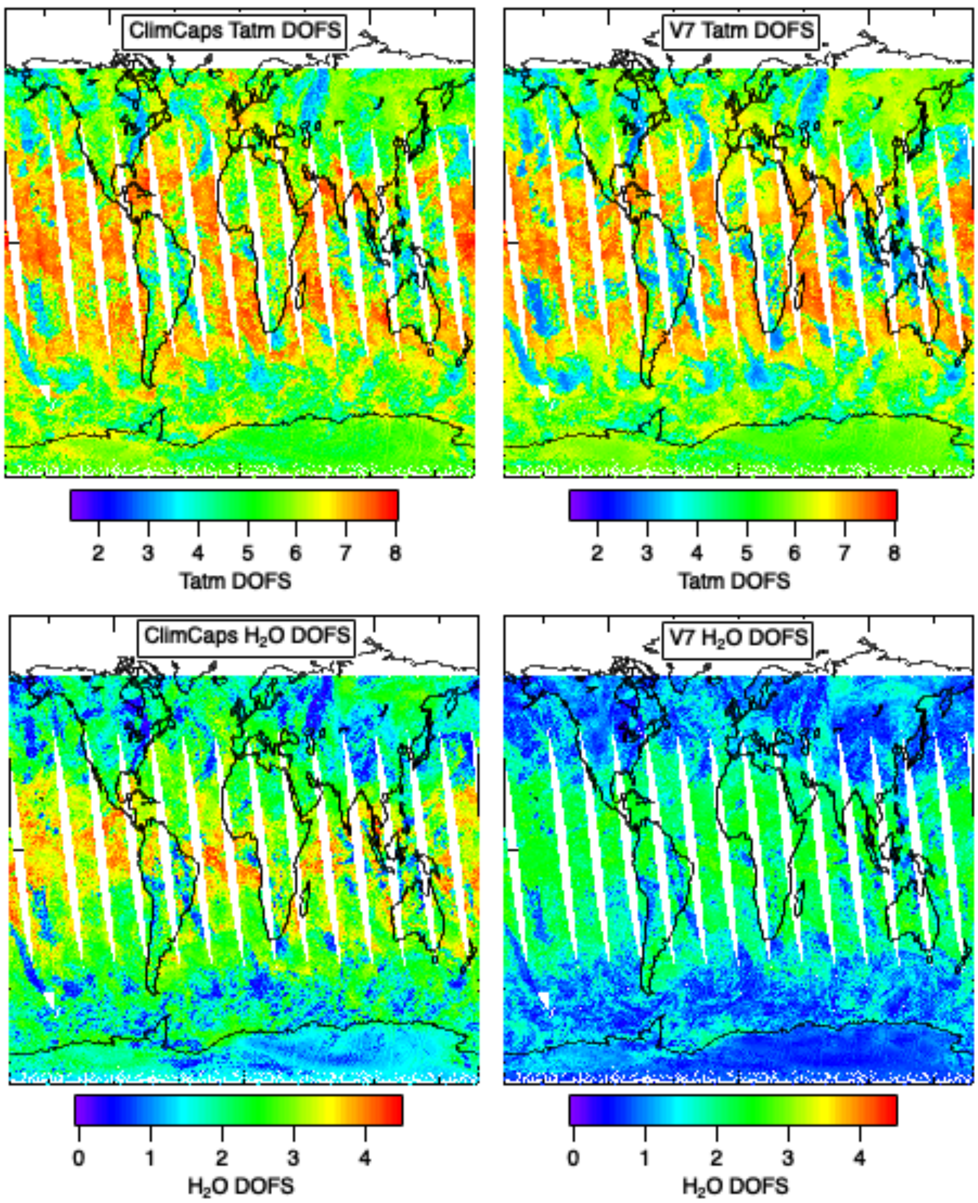


Figure 9.2.1 Sample maps of degree of freedom (DOF) calculated from temperature (top row) and water vapor (bottom row) averaging kernels for CLIMCAPS-*Aqua* retrievals (left column) and AIRS V7 retrievals (right column). Data are from AIRS daytime observations of January 1, 2011 and use IR+MW retrievals.

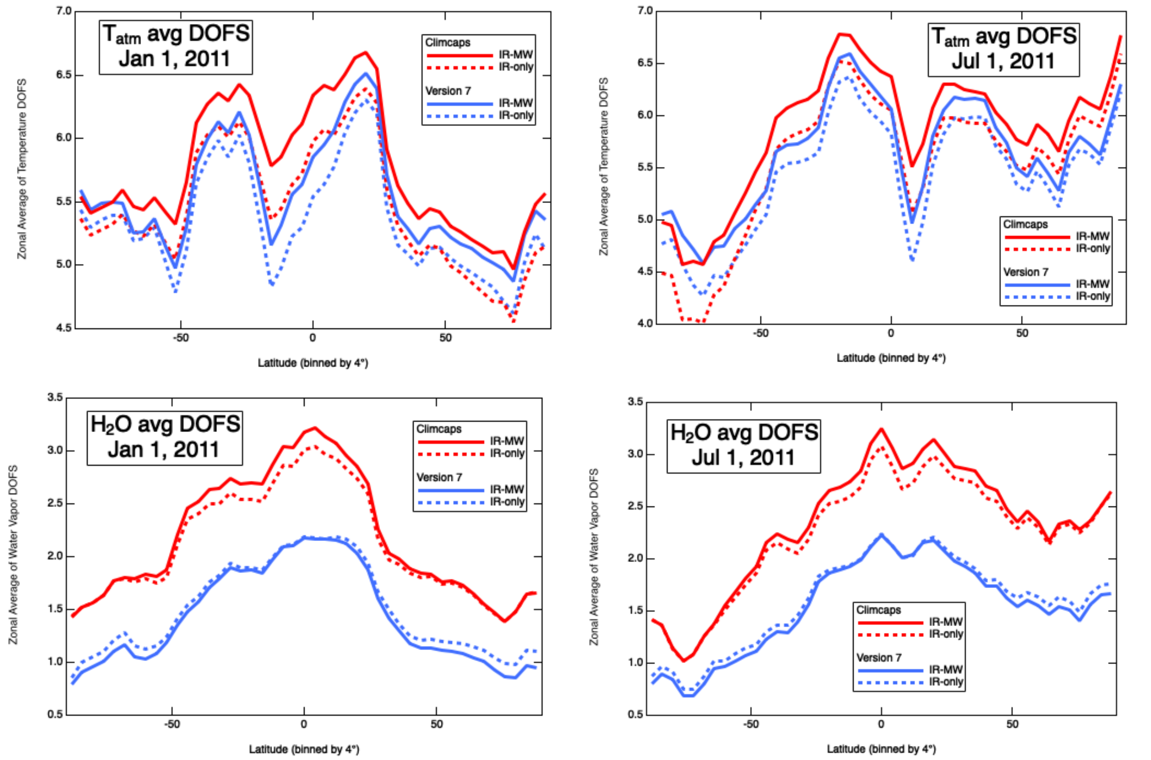


Figure 9.2.2. Zonal averages (binned by 4° latitude) of DOF for temperature (top row) and water vapor (bottom row) for the full days of January 1, 2011 (left column) and July 1, 2011 (right column).

Reference

Boylan, P., Wang, J., Cohn, S. A., Fetzer, E., Maddy, E. S., and Wong, S. (2015), Validation of AIRS version 6 temperature profiles and surface-based inversions over Antarctica using Concordiasi dropsonde data, *J. Geophys. Res. Atmos.*, 120, 992– 1007, doi:10.1002/2014JD022551.

Durre, I., R. S. Vose, and D. B. Wuertz (2006), Overview of the Integrated Global Radiosonde Archive, *J. Climate*, *19*, 53-68.

Fetzer, E. J., B. H. Lambrigtsen, A. Eldering, H. H. Aumann, and M. T. Chahine (2006), Biases in total precipitable water vapor climatologies from Atmospheric Infrared Sounder and Advanced Microwave Scanning Radiometer, *J. Geophys. Res., 111*, D09S16, doi:10.1029/2005JD006598.

Granskog, M. A., P. Assmy, S. Gerland, G. Spreen, H. Steen and L. H. Smedsrud, 2016: Arctic research on thin ice: Consequences of Arctic sea ice loss, Eos, 97, doi:10.1029/2016EO044097.

Maddy, E. S. and Barnet, C. D., Vertical resolution estimates in Version 5 of AIRS operational retrievals, IEEE T. Geosci. Remote, 46, 2375–2384, 2008.

Milstein, A. B., and W. J. Blackwell, Neural network temperature and moisture retrieval algorithm validation for AIRS/AMSU and CrIS/ATMS, J. Geophys. Res. Atmos., 121, 1414–1430, doi:10.1002/2015JD024008, 2016.

Monarrez, R., ed., 2020, NASAS-NPP and NOAA-20 (JPSS-1) CLIMCAPS CrIS and ATMS Level-2 Products User Guide: File Format and Definition, available at: https://docserver.gesdisc.eosdis.nasa.gov/public/project/Sounder/CLIMCAPS.V2.README.pdf.

Kahn, B. H., Chahine, M. T., Stephens, G. L., Mace, G. G., Marchand, R. T., Wang, Z., Barnet, C. D., Eldering, A., Holz, R. E., Kuehn, R. E., and Vane, D. G.: Cloud type comparisons of AIRS, CloudSat, and CALIPSO cloud height and amount, Atmos. Chem. Phys., 8, 1231–1248, https://doi.org/10.5194/acp-8-1231-2008, 2008.

Kahn, B. H., Irion, F. W., Dang, V. T., Manning, E. M., Nasiri, S. L., Naud, C. M., Blaisdell, J. M., Schreier, M. M., Yue, Q., Bowman, K. W., Fetzer, E. J., Hulley, G. C., Liou, K. N., Lubin, D., Ou, S. C., Susskind, J., Takano, Y., Tian, B., and Worden, J. R.: The Atmospheric Infrared Sounder version 6 cloud products, Atmos. Chem. Phys., 14, 399–426, https://doi.org/10.5194/acp-14-399-2014, 2014.

Kalmus, P., S. Wong, and J. Teixeira (2015), The Pacific subtropical cloud transition: a MAGIC assessment of AIRS and ECMWF profiles in the northeast Pacific, *IEEE Geosci. And Remote Sens. Lett.*, 12, 1586-1590.

OMI Team, Ozone Monitoring Instrument (OMI) Data User's Guide, 2012. Available at https://docserver.gesdisc.eosdis.nasa.gov/repository/Mission/OMI/3.3_ScienceDataProductDocumentation/3.3.2_ProductRequirements_Designs/README.OMI_DUG.pdf

Rodgers, C. D.: Inverse Methods for Atmospheric Sounding: Theory and Practice, World Scientific Publishing, Singapore, 2000.

Roman, J., Knuteson, R., August, T., Hultberg, T., Ackerman, S., and Revercomb, H. (2016), A global assessment of NASA AIRS v6 and EUMETSAT IASI v6 precipitable water vapor using ground-based GPS SuomiNet stations, *J. Geophys. Res. Atmos.*, 121, 8925– 8948, doi:10.1002/2016JD024806.

Roman, J., et al., 2021, Global Evaluation of CLIMCAPS-*Aqua* Temperature and Humidity Profiles to ECMWF, in preparation.

Roman, J., et al., 2021, Evaluation of CLIMCAPS-*Aqua* Total Column Water Vapor to SuomiNet GPS Stations and AMSR-E observations, in preparation.

Smith, N.; Barnet, C.D. 2019, Uncertainty Characterization and Propagation in the Community Long-Term Infrared Microwave Combined Atmospheric Product System (CLIMCAPS). *Remote Sens.* **2019**, *11*, 1227.

Smith, N. and Barnet, C. D.: CLIMCAPS observing capability for temperature, moisture, and trace gases from AIRS/AMSU and CrIS/ATMS, Atmos. Meas. Tech., 13, 4437–4459, https://doi.org/10.5194/amt-13-4437-2020, 2020.

Sotiropoulou, G., M. Tjernström, J. Sedlar, P. Achtert, B. J. Brooks, I. M. Brooks, P. O. G. Persson, J. Prytherch, D. J. Salisbury, M. D. Shupe, P. E. Johnston and D. Wolfe, 2016: Atmospheric Conditions during the Arctic Clouds in Summer Experiment (ACSE): Contrasting Open Water and Sea Ice Surfaces during Melt and Freeze-Up Seasons, J. Clim., 29, 8721-8744, doi:10.1175/JCLI-D-16-0211.1.

Susskind, J., Barnet, C. D., and Blaisdell, J., Retrieval of atmospheric and surface parameters from AIRS/AMSU/HSB data in the presence of clouds, IEEE Trans. Geosci. Remote Sens., 41(2), 390–409, 2003.

Tian, B., & Hearty, T. (2020). Estimating and removing the sampling biases of the AIRS Obs4MIPs V2 data. *Earth and Space Science*, 7, e2020EA001438. https://doi.org/10.1029/2020EA001438

Tjernström, M. and coauthors, 2014: The Arctic Summer Cloud Ocean Study (ASCOS): overview and experimental design, Atmos. Chem. Phys., 14, 2823-2869, doi:10.5194/acp-14-2823-2014.

Veefkind, J. P., de Haan, J. F., Brinksma, E. J., Kroon, M., and Levelt, P. F., Total Ozone from the Ozone Monitoring Instrument (OMI) Using the DOAS Technique, IEEE Trans. Geo. Rem. Sens., 44(5), 1239-1244, 2006.

Wang, T., J. Roman, Q. Yue, and S. Wong, ed., 2020, TestReport of Performance of CLIMCAPS-SNPPand CLIMCAPS-JPSS1Retrievals, available at: https://docserver.gesdisc.eosdis.nasa.gov/public/project/Sounder/CLIMCAPS.V2.Test.Report.pdf

Wang. T., et al., 2021, Validation of AIRS Temperature and Water Vapor Retrievals using the IGRA radiosonde observations, in preparation.

Wong, S., E. J. Fetzer, M. Schreier, G. Manipon, E. F. Fishbein, B. H, Kahn, Q. Yue, and F. W. Irion (2015), Cloud-induced uncertainties in AIRS and ECMWF temperature and specific humidity, *J. Geophys. Res.*, 120, doi:10.1002/2014JD022440.

Yue, Q., E. J. Fetzer, B. H. Kahn, S. Wong, G. Manipon, A. Guilaume, and B. Wilson, 2013: Cloud-state-dependent sampling in AIRS observations based on CloudSat cloud classification, *J. of Climate*, 26, 8357–8377, doi: http://dx.doi.org/10.1175/JCLI-D-13-00065.1.

Yue, Q., E. J. Fetzer, L. Wang, I. Tkatcheva, B. H. Kahn, S. Wong, M. Schreier, 2021: The consistency of cloud data records from the simultaneous observations by *Aqua* and SNPP, in preparation.

Yue, Q., T. Wang, S. Wong, E. Manning, E. J. Fetzer, B. Lambrigsten, 2021: Validation of AIRS temperature and water vapor profiles in the polar region, in preparation.

Yue, Q., and B. Lambrigtsen, ed., 2017, AIRS V6 Test Report Supplement: Performance of AIRS+AMSU vs. AIRS-only Retrievals, available at: https://docserver.gesdisc.eosdis.nasa.gov/repository/Mission/AIRS/3.3_ScienceDataProduct

Documentation/3.3.5_ProductQuality/V6_Test_Report_Supplement_Performance_of_AIRS+AMSU_vs_AIRS-Only_Retrievals.pdf.
Yue, Q., and B. Lambrigtsen, ed., 2019, AIRS V7 L2 Performance Test and Validation Report, available at: https://docserver.gesdisc.eosdis.nasa.gov/public/project/AIRS/V7_L2_Performance_Test_and_Validation_report.pdf